

Thu Feb 10 14:37:07 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:46:53 ; Search time 40 Seconds
(without alignments)
921.276 Million cell updates/sec

Title: US-10-056-405-2
Perfect score: 383
Sequence: 1 MYNFKVFKKCAPNGKVTLY.....ADVFPRQDTIDQASVDFE 383

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: PIR.79:*
2: p1r1:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	5.2	364	2 A34867	arrestin - fruit f
2	17	4.4	363	2 A55081	arrestin 1 - blueb
3	16	4.2	381	2 B56607	arrestin homolog -
4	12	3.1	407	2 A56607	arrestin homolog -
5	11	2.9	401	2 A34856	49K photoreceptor
6	11	2.9	401	2 B15566	arrestin homolog -
7	11	2.9	401	2 B55081	arrestin 2 - blueb
8	11	2.9	405	2 S68254	arrestin isoform 2
9	11	2.9	407	2 S68253	arrestin isoform 1
10	11	2.9	409	2 S18984	arrestin - human (
11	11	2.9	410	2 A59279	beta-arrestin 2 -
12	11	2.9	415	2 S68255	arrestin isoform 1
13	11	2.9	420	2 A47140	arrestin arf3L - b
14	9	2.3	387	2 I55423	arrestin-C - human
15	9	2.3	388	2 S38943	arrestin - human
16	9	2.3	1773	2 T05128	hypothetical prote
17	8	2.1	403	2 J80066	S-antigen - mouse
18	8	2.1	403	2 S03960	S-antigen - rat
19	8	2.1	404	2 A28404	S-antigen - bovine
20	8	2.1	405	2 A30357	retinal S-antigen
21	8	2.1	412	2 F87709	DNA/pantochinase m
22	8	2.1	418	2 B46682	beta-arrestin 1, s
23	8	2.1	418	2 A34851	beta-arrestin, bra
24	8	2.1	418	2 B43404	beta-arrestin1 - r
25	8	2.1	448	2 D71852	probable signal re
26	8	2.1	448	2 H64663	signal recognition
27	7	1.8	53	2 AF2764	nitrogen fixation
28	7	1.8	57	2 P70202	protein-lysine k
29	7	1.8	83	1 MSWLSH	ES protein - human

30	7	1.8	92	2 B82603	hypothetical prote
31	7	1.8	111	2 S10222	hypothetical prote
32	7	1.8	139	2 E97426	flagellar basal-bo
33	7	1.8	139	2 AE2544	flagellar basal-bo
34	7	1.8	142	2 F81048	conserved hypochet
35	7	1.8	145	2 C64447	hypothetical prote
36	7	1.8	149	2 A87346	hypothetical prote
37	7	1.8	151	2 A83639	osmotically induc
38	7	1.8	155	2 C81827	hypothetical prote
39	7	1.8	161	2 A27873	aliphococyanin al
40	7	1.8	166	2 A95969	hypothetical prote
41	7	1.8	174	2 S39877	carQ protein - Myx
42	7	1.8	216	2 A61259	gycoprotein S - p
43	7	1.8	217	2 B64080	hypothetical prote
44	7	1.8	231	2 S73469	probable lipoprote
45	7	1.8	231	2 A82084	MTR/SAH nucleoside

ALIGNMENTS

* RESULT 1
A34867
arrestin - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 09-Jul-2004
C/Accession: A34867; B34868
R/Smith, D.P.; Shieh, B.H.; Zuker, C.S.
Proc. Natl. Acad. Sci. U.S.A. 87, 1003-1007, 1990
A/Title: Isolation and structure of an arrestin gene from Drosophila.
A/Reference number: A34867; PMID:90138925; PMID:1689056
A/Accession: A34867
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-364 <SMI>
A/Cross-references: UNIPROT:P15372; GB:M30177
R/Hyde, D.R.; Mecklenburg, K.L.; Pollock, J.A.; Vitellie, T.S.; Benzer, S.
Proc. Natl. Acad. Sci. U.S.A. 87, 1008-1012, 1990
A/Title: Twenty Drosophila visual system cDNA clones: one is a homolog of human arrestin
A/Reference number: A34868; PMID:90138946; PMID:2105491
A/Accession: A34868
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-364 <HYD>
A/Cross-references: GB:M30140; NID:G156961; PIDN:AAA28380.1; PID:G156962
C/Genetics:
A/Gene: FlyBase:Ar1
A/Cross-references: FlyBase:FBgn0000120
A/Intons: 36/3; 306/3; 349/3
C/Superfamily: arrestin

Query Match 5.2%; Score 20; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 38-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 VKKIKAMVOGQVDFPONG 247
Db 227 VKKIKAMVOGQVDFPONG 246

RESULT 2
A55081
arrestin 1 - bluebottle fly (Calliphora vicina)
C/Species: Calliphora vicina
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C/Accession: A55081; S44291
R/Planger, A.; Malicki, D.; Whitney, M.; Paulsen, R.
J. Biol. Chem. 269, 26869-26975, 1994
A/Title: Mechanism of arrestin 2 function in rhodameric photoreceptors.
A/Reference number: A55081; PMID:95014564; PMID:7929436
A/Accession: A55081
A/Status: preliminary
A/Molecule type: mRNA

This Page Blank (uspto)

This Page Blank (uspto)

Local Similarity 100.0%; DB 2; Length 363;
Matches 17; Conservative 0; Mismatches 0;

RESULT 3
R56607

Query Match

Best Local Similarity 100.0%; score 16; DB 2; Length 381;
Matches 16; Conservative 0.0; Pred. No. 4.3e-08;

RESULT 4

A56607
 arrestin homolog - migratory locust
 C:Species: locusta migratoria (migratory locust)
 C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #next_change 09-Jul-2004
 R:Accession: A56607
 R:Author: R. Remington, K. Freidberg, J. Krieger, T. Breer, H. Well
 R:Title: Arrestin: A56607

```

\Reference number: subtypes in insect antennae.
\Contents: A56607; MUID:93199955; PMID:8452755
\Accession: A56607
\Status: Preliminary
\Molecule type: mRNA
\Residues: 1-407 <RAM>
\Cross-references: UNIPROT:P3122; GB:S57174; NID:G298755;
Note: sequence extracted from NCBI backbone (NCBIN:127933,
Superfamily: arrestin

```

Query Match

```
Oy      62 RYGRREDEVMGL 73          | Indels    0; Gaps    0;
Db       69 RYGRREDEVMGL 80          | Mismatches 0; Conservatve 0.00063; Pred.No. 0.0%; Score 12; DB 2; Length 407;
                        Matches 12; Local Similarity 100.0%;
```

RESULT 5

49K photoreceptor protein - fruit fly (*Drosophila melanogaster*)

Supplementally: arrestin

```

Query Match      2.9%; Score 11; DB 2; Length 401,
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 11; Conservative 0

```

	Mismatches	Indels	Gaps
QY 62 RYGRDEEVDWG 72 	0;	0;	0;
Dd 61 RYGRDEEVDWG 71			

RESULT 6

s11566
 arrestin homolog - fruit fly (Drosophila miranda)
 C/Species: Drosophila miranda
 C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 R/Krishnan, R.; Ganguly, R.
 Nucleic Acids Res. 18, 5894, 1990
 A>Title: Nucleotide sequence of the arrestin-like 49 kD protein gene of Drosophila miranda
 A/Accession: S11566, MUID:91016944, PMID:2216789
 A>Status: preliminary; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-401 <R1>
 A/Cross-references: UNIPROT:P19108, EMBL:X54084, NID:G7612, PIDN:CA438019.1, PID:9295739
 C/Genetics:
 A/Genes: FlyBase:Dmir/Arr2
 A/Cross-references: FlyBase:FBgn0012552
 C/Introns: 286/1, 351/1
 C/Superfamily: arrestin

Query Match

Best Local Similarity	2.9%;	Score 11;	DB 2;	Length 401;
Matches 11;	Conservative 0;	Mismatches 0;		

	Indels	Gaps
QY 62 RYGREDEYWG 72	0	0
DB 61 RYGREDEYWG 71	0	0

RESULT 7

B55081
arrestin 2 - bluebottle fly (*Calliphora vicina*)
C/Species: *Calliphora vicina*
C/Date: 08-Jul-1995 #sequence
Accession: B55081, S44292
R/Planger: A.; Mallick, D.; Whitney, M.; Paulsen, R.
J. Biol. Chem. 269, 26960-26975, 1994
Article: Mechanism of arrestin 2 function in rhodameric photoreceptors.
A/Accession number: A55081; MVID:95014564; PMID:7929436
Status: draft

A;Molecule type: MR

A; residues: 1-401 <PLA>
A; Cross-references: UNIPROT:P51487; EMBL:X79073; NID:g483583; PIDN:CA55673.1; PID:g43355
A; Note: the source is designated as *Calliphora erythrocephala*
C; Superfamily: arrestin

Query Match 2.9%; Score 11; DB 2; Length 401;
A: Molecule type: mRNA
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 RYGRDEDEVMG 72
|||||
Db 61 RYGRDEDEVMG 71

RESULT 8
S68254
arrestin isoform 2S, erythrocyte - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S68254
R:Jahn, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motaïs, R.; Fievet, B.
Biochem. J. 316, 497-506, 1996
A:Title: Trout red blood cell arrestin (TRCarr), a novel member of the arrestin family:
A:Reference number: S68253; PMID:8687393
A:Accession: S68254
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-405 <JAH>
A:Cross-references: UNIPROT:P51467; GB:U48410
C:Superfamily: arrestin

Query Match 2.9%; Score 11; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 0.0066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELFPVLMHPRK 351
|||||
Db 339 ELFPVLMHPRK 349

RESULT 9
S68253
arrestin isoform 1S, erythrocyte - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S68253
R:Jahn, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motaïs, R.; Fievet, B.
Biochem. J. 316, 497-506, 1996
A:Title: Trout red blood cell arrestin (TRCarr), a novel member of the arrestin family:
A:Reference number: S68253; PMID:8687393
A:Accession: S68253
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-407 <JAH>
A:Cross-references: UNIPROT:P51466; GB:U48410; NID:G1215723; PIDN:AA16954.1; PID:G12157
C:Superfamily: arrestin

Query Match 2.9%; Score 11; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELFPVLMHPRK 351
|||||
Db 339 ELFPVLMHPRK 349

RESULT 10
S18984
arrestin - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S18984
R:Rapoport, B.
Submitted to the EMBL Data Library, November 1991
A:Description: Isolation of a novel cDNA belonging to the arrestin family from human tly
A:Reference number: S18984
A:Accession: S18984

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-409 <RAP>
A:Cross-references: UNIPROT:P32121; EMBL:Z11501; NID:G28850; PIDN:CAA7577.1; PID:G28851
C:Superfamily: arrestin

Query Match 2.9%; Score 11; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.0069;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELFPVLMHPRK 351
|||||
Db 339 ELFPVLMHPRK 349

RESULT 11
A59279
beta-arrestin 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: A59279; A43404; I70112
R:Attramadal, H.; Arriza, J.L.; Aoki, C.; Dawson, T.M.; Codina, J.; Kwatra, M.M.; Snyder
Submitted to GenBank, August 1995
A:Reference number: A59279
A:Accession: A59279
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-410 <ATT>
A:Cross-references: UNIPROT:P29067; GB:M91590; NID:G949986; PIDN:AAA74460.1; PID:G203104
A:Experimental source: strein Sprague-Dawley; stage adult; tissue type brain; tissue 1lb
R:Attramadal, H.; Arriza, J.L.; Aoki, C.; Dawson, T.M.; Codina, J.; Kwatra, M.M.; Snyder
J. Biol. Chem. 267, 17882-17890, 1992
A:Title: Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene family.
A:Reference number: A43404; PMID:92388146; PMID:1517224
A:Accession: A43404
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-320, 'A', 322-325, 'A', 327-330, 'A', 332-410 <ATT>
A:Cross-references: GB:M91590; NID:G949986
A:Note: sequence extracted from NCBI backbone (NCBI:P:112790)
R:Craef, C.M.; Whitmore, D.H.; Wiedemann, A.F.
J. Biol. Chem. 269, 4613-4619, 1994
A:Title: Cone arrestin identified by targeting expression of a functional family.
A:Reference number: I55423; PMID:94140898; PMID:8308033
A:Accession: I70112
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 295-410 <CRA>
A:Cross-references: EMBL:U03627; NID:G458202; PIDN:AAA17551.1; PID:G458203
C:Superfamily: arrestin

Query Match 2.9%; Score 11; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELFPVLMHPRK 351
|||||
Db 340 ELFPVLMHPRK 350

RESULT 12
S68255
arrestin isoform 1L, erythrocyte - rainbow trout
C:Species: Oncorhynchus mykiss (rainbow trout)
C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S68255
R:Jahn, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motaïs, R.; Fievet, B.
Biochem. J. 316, 497-506, 1996
A:Title: Trout red blood cell arrestin (TRCarr), a novel member of the arrestin family:
A:Reference number: S68253; PMID:8687393
A:Accession: S68255
A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-415 <JAH>
A:Cross-references: UNIPROT:P51468; GB:U48410
C:Superfamily: arrestin

Query Match
Best Local Similarity 2.9%; Score 11; DB 2; Length 415;
Matches 11; Conservative 100.0%; Pred. No. 0.0069; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELPPVLMHPKP 351
DB 347 ELPPVLMHPKP 357

RESULT 13

A47140

arrestin ar33l - bovine

N:Contains: arrestin ar33

C:Species: Bos primigenius taurus (cattle)

C>Date: 16-Feb-1994 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004

R:Accession: A47140; B47140; J02051

J. Biol. Chem. 269, 15640-15648, 1993

A:Title: Polypeptide variants of beta-arrestin and arrestin3.

A:Reference number: A47140; MUID:93340166; PMID:8340388

A:Accession: A47140

A:Molecule type: mRNA

A:Residues: 1-420 <STE>

A:Cross-references: UNIPROT:P32120; GB:L14641

A:Experimental source: brain

A:Note: sequence extracted from NCBI backbone (NCBIN:136006, NCBI:P:136007)

A:Accession: B47140

A:Molecule type: mRNA

A:Residues: 1-362,374-420 <ST2>

C:Comment: This protein plays a role in the regulation of G-protein-coupled receptors.

C:Superfamily: arrestin

F:1-420/Product: arrestin

F:1-362,374-420/Product: arrestin ar33l #status predicted <MAT>

Query Match

Best Local Similarity 2.9%; Score 11; DB 2; Length 420;

Matches 11; Conservative 100.0%; Pred. No. 0.007; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELPPVLMHPKP 351
DB 339 ELPPVLMHPKP 349

RESULT 14

I55423

arrestin-C - human

C:Species: Homo sapiens (man)

C>Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 13-Sep-1998

R:Accession: I55423

J. Biol. Chem. 269, 4613-4619, 1994

A:Title: Cone arrestin identified by targeting expression of a functional family.

A:Reference number: I55423; MUID:94140898; PMID:8308033

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-387 <RES>

A:Cross-references: EMBL:U03626; NID:9458200; PID:9458201

C:Superfamily: arrestin

Query Match
Best Local Similarity 2.3%; Score 9; DB 2; Length 387;
Matches 9; Conservative 100.0%; Pred. No. 0.076; Mismatches 0; Indels 0; Gaps 0;

QY 22 GKRDVVDHV 30
DB 19 GKRDVVDHV 27

RESULT 15

S38943

arrestin - human

C:Species: Homo sapiens (man)

C>Date: 19-May-1994 #sequence_revision 17-Nov-1995 #text_change 09-Jul-2004

R:Accession: S38943

R:Murakami, A.; Yajima, T.; Sakuma, H.; McLaren, M.J.; Inana, G.

FEBS Lett. 334, 203-209, 1993

A:Title: X-arrestin: a new retinal arrestin mapping to the X chromosome.

A:Reference number: S38943; MUID:94039835; PMID:8224247

A:Accession: S38943

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-388 <MUR>

A:Cross-references: UNIPROT:P36575; GB:S66793; NID:9439830; PID:AA828824.1; PID:9439831

C:Superfamily: arrestin

Query Match

Best Local Similarity 2.3%; Score 9; DB 2; Length 388;

Matches 9; Conservative 100.0%; Pred. No. 0.076; Mismatches 0; Indels 0; Gaps 0;

QY 22 GKRDVVDHV 30
DB 19 GKRDVVDHV 27

Search completed: February 10, 2005, 00:02:18
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:39:27 ; Search time 41 Seconds
(without alignments)
898.805 Million cell updates/sec

Title: US-10-056-405-2
Perfect score: 1973
Sequence: 1 MVYFKVFKKCAPNGKVTLY.....ADVETFRQDTIDQASVDFE 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499	76.0	381	2	arrestin homolog -
2	1384	70.1	364	2	arrestin - fruit f
3	1335	67.7	363	2	arrestin 1 - blueb
4	993.5	50.4	407	2	arrestin homolog -
5	977.5	49.5	401	2	arrestin homolog -
6	975.5	49.4	401	2	49k photoreceptor
7	942.5	47.8	401	2	arrestin 2 - blueb
8	903.5	45.8	418	2	beta-arrestin, bra
9	903.5	45.8	418	2	B43404
10	891.5	45.2	418	2	beta-arrestin 1, r
11	887	45.0	420	2	arrestin ar3L - b
12	882	44.7	409	2	arrestin - human (
13	871	44.1	407	2	arrestin isoform 1
14	869.5	44.1	410	2	beta-arrestin 2 -
15	865	43.8	454	2	hypothetical prote
16	863	43.7	415	2	arrestin isoform 1
17	861	43.6	405	2	arrestin isoform 2
18	810	39.1	387	2	cone arrestin - Af
19	785.5	39.8	389	2	cone arrestin - hu
20	784.5	39.8	389	2	cone arrestin - no
21	783.5	39.7	403	2	S-antigen - mouse
22	780.5	39.6	388	2	S-antigen - human
23	775.5	39.3	403	2	S-antigen - bovine
24	768.5	39.0	404	2	retinal S-antigen
25	761	38.6	405	2	rod arrestin - nor
26	731.5	37.1	396	2	rod arrestin - C - human
27	731	37.0	396	2	rod arrestin - bul
28	729.5	37.0	396	2	C-arrestin - rat (
29	134	6.8	92	2	170113

30	129	6.5	378	2	T28112	hypothetical prote
31	118.5	6.0	427	2	T23954	hypothetical prote
32	111.5	5.7	364	2	T26507	hypothetical prote
33	111.5	5.7	471	2	T15155	hypothetical prote
34	111.5	5.7	2413	2	S34670	splicing factor PR
35	110.5	5.6	344	2	T33421	hypothetical prote
36	110	5.6	456	2	T29948	hypothetical prote
37	107.5	5.4	312	2	T20949	hypothetical prote
38	107.5	5.4	460	2	T23955	hypothetical prote
39	106.5	5.4	460	2	T33420	hypothetical prote
40	106	5.4	347	2	S31144	mk1 protein - Myco
41	104	5.3	338	2	T87145	ABC transporter AT
42	102.5	5.2	464	2	T16826	hypothetical prote
43	101.5	5.1	1804	2	T34518	neatin - golden ha
44	100	5.1	359	2	B70534	probable ribonucle
45	99.5	5.0	600	2	S65788	ku antigen 70K cha

ALIGNMENTS

RESULT 1
B56607
arrestin homolog - tobacco budworm
C:Species: Heliothis virescens (tobacco budworm)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C:Accession: B56607
R:Ramirez, K.; Freitag, J.; Krieger, J.; Breer, H.
Cell. Signal. 5, 69-80, 1993
A:Title: Arrestin-subtypes in insect antennae.
A:Reference number: A56607; MUID:9319955; PMID:8452755
A:Accession: B56607
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-381 <RAM>
A:Cross-references: UNIPROT:P55274
A:Experimental source: antennae
A>Note: sequence extracted from NCBI backbone (NCBIP:127926)
C:Superfamily: arrestin

Query Match	76.0%; Score 1499; DB 2; Length 381;
Best Local Similarity	72.7%; Pred. No. 7.7e-112;
Matches 277; Conservative 51; Mismatches 51; Indels 2; Gaps 2;	
Qy	1 MVYFKVFKKCAPNGKVTLYMGKRDVYHSGVPEIDGIYVLDDEYIRDNKRVGOIVCS 60
Db	1 MVYFKVFKKCAPNGKVTLYMGKRDVYHSGVPEIDGIYVLDDEYIRDNKRVGOIVCS 59
Qy	61 FRYGREDEVMGILNFOKELCLASEQIYPRPEKSDKEOTKLOERLLKXLSNAIPFTFNIS 120
Db	60 FRYGREDEVMGILNFOKELCLASEQIYPRPEKSDKEOTKLOERLLKXLSNAIPFTFNIS 119
Qy	121 PNAHSVTLQOGEEDNDPCGVSYVYKIPAGESEFTDTHRSYVTLGIRKIQAPPTQOGQ 180
Db	120 PNAHSVTLQOGEEDNDPCGVSYVYKIPAGESEFTDTHRSYVTLGIRKIQAPPTQOGQ 179
Qy	181 OPTLVKRDPMLSGELLEVTLDKOLYHGERGVNICTRNNSKNVKKIKAVVQGVGD 240
Db	180 OPTLVKRDPMLSGELLEVTLDKOLYHGERGVNICTRNNSKNVKKIKAVVQGVGD 239
Qy	241 VVLFQNGSYRNTVSLTSEGCPIOPGSSLOKWVYLPPLSSNQKRGIALDGOIKRQDQ 300
Db	240 VVLFQNGSYRNTVSLTSEGCPIOPGSSLOKWVYLPPLSSNQKRGIALDGOIKRQDQ 299
Qy	301 CLASTLLAOPDQDAFGVILSYAVKVLPLGALGSLASLPLVLMHPKGTAKVIA 360
Db	300 CLASTLLAOPDQDAFGVILSYAVKVLPLGALGSLASLPLVLMHPKGTAKVIA 358
Qy	361 DSQADVETFRQDTIDQASVD 381
Db	359 DSQADVETFRQDTIDQASVD 379

RESULT 2

A34867

arrestin - fruit fly (*Drosophila melanogaster*)C/Species: *Drosophila melanogaster*

C/Date: 20-Jul-1990 #sequence

C/Accession: A34867; A34868

R/Smith, D.P.; Shieh, B.H.; Zuker, C.S.

Proc. Natl. Acad. Sci. U.S.A. 87, 1003-1007, 1990

A/Title: Isolation and structure of an arrestin gene from *Drosophila*.

A/Reference number: A34867; MUID:90138925; PMID:1689056

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-364 <SWT>

A/Cross-references: UNIPROT:P15372; GB:M30177

Proc. Natl. Acad. Sci. U.S.A. 87, 1008-1012, 1990

A/Title: Twenty *Drosophila* visual system cDNA clones: one is a homolog of human arrestin

A/Reference number: A34868; MUID:90138926; PMID:2105491

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-364 <HYD>

A/Cross-references: GB:M30140; NID:9156961; PIDN:AAA28380.1; PID:9156962

A/Genetic: FlyBase:Arri1

A/Introns: 36/3; 306/3; 349/3

C/Superfamily: arrestin

Query Match

Best Local Similarity 70.1%; Score 1384; DB 2; Length 364;

Matches 238; Conservative 51; Mismatches 54; Indels 4; Gaps 2;

QY 1 MVTNFKVKKCAPNGKVTLYMGRDPVDHVGVEPIDGIIVLDEYTRDNKRYFGQIVGS 60

DB 1 MVTNFKVKKCAPNGKVTLYMGRDPVDHVGVEPIDGIIVLDEYTRDNKRYFGQIVGS 60

QY 61 FRYGRDEDMIGRFOKELELVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

DB 61 FRYGRDEDMIGRFOKELELVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

QY 121 PNASSVTLQOGEEDNDGPGCVSYVYKIFAGESEDTTRHRSSTVTLGIRKIQAPTKOQ 119

DB 121 PNASSVTLQOGEEDNDGPGCVSYVYKIFAGESEDTTRHRSSTVTLGIRKIQAPTKOQ 119

QY 120 PSSPASVTLQOQKASDESQPCGVQFVAVFAGENDCDRSHRSSTVTLGIRKIQAPTKOQ 180

DB 120 PSSPASVTLQOQKASDESQPCGVQFVAVFAGENDCDRSHRSSTVTLGIRKIQAPTKOQ 180

QY 181 OPTLVKRDPMFLSPGELLEVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

DB 181 OPTLVKRDPMFLSPGELLEVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

QY 241 VVLFQNGSVYRTVYASLETSEGGCPIDPGSSLOKVMYLPPLISNKKORRGIALDQIRKQD 239

DB 241 VVLFQNGSVYRTVYASLETSEGGCPIDPGSSLOKVMYLPPLISNKKORRGIALDQIRKQD 239

QY 301 CLASTTLTAQPDORDAFVGIISYAVKVLFLGALGSELPAELPFVLMHPKPTAKVIHA 360

DB 301 CLASTTLTAQPDORDAFVGIISYAVKVLFLGALGSELPAELPFVLMHPKPTAKVIHA 360

QY 361 DSQADVE 367

DB 361 DSQADVE 367

QY 357 EAEGSIE 363

DB 357 EAEGSIE 363

RESULT 3

A55081

arrestin 1 - bluebottle fly (*Calliphora vicina*)C/Species: *Calliphora vicina*

C/Date: 08-Jul-1995 #sequence

C/Accession: A55081; A55082

R/Plangger, A.; Malicki, D.; Whitney, M.; Paulsen, R.

J. Biol. Chem. 269, 26969-26975, 1994

A/Title: Mechanism of arrestin 2 function in rhodometric photocoreceptors.

A/Reference number: A55081; MUID:95014564; PMID:7929436

A/Accession: A55081

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-363 <PLA>

A/Cross-references: UNIPROT:P15486; GB:X79072; NID:9483527; PIDN:CAA55672.1; PID 94

A/Note: the source is designated as *Calliphora erythrocephala*

C/Superfamily: arrestin

Query Match

Best Local Similarity 67.7%; Score 1335; DB 2; Length 363;

Matches 248; Conservative 57; Mismatches 55; Indels 8; Gaps 2;

QY 1 MVTNFKVKKCAPNGKVTLYMGRDPVDHVGVEPIDGIIVLDEYTRDNKRYFGQIVGS 60

DB 1 MVTNFKVKKCAPNGKVTLYMGRDPVDHVGVEPIDGIIVLDEYTRDNKRYFGQIVGS 60

QY 61 FRYGRDEDMIGRFOKELELVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

DB 61 FRYGRDEDMIGRFOKELELVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

QY 121 PNASSVTLQOGEEDNDGPGCVSYVYKIFAGESEDTTRHRSSTVTLGIRKIQAPTKOQ 119

DB 121 PNASSVTLQOGEEDNDGPGCVSYVYKIFAGESEDTTRHRSSTVTLGIRKIQAPTKOQ 119

QY 120 PSSPASVTLQOQKASDESQPCGVQFVAVFAGENDCDRSHRSSTVTLGIRKIQAPTKOQ 180

DB 120 PSSPASVTLQOQKASDESQPCGVQFVAVFAGENDCDRSHRSSTVTLGIRKIQAPTKOQ 180

QY 181 OPTLVKRDPMFLSPGELLEVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

DB 181 OPTLVKRDPMFLSPGELLEVTLDKQLYHGERIGVNICIRNNSKNKVKIKKAMVOQGYD 240

QY 241 VVLFQNGSVYRTVYASLETSEGGCPIDPGSSLOKVMYLPPLISNKKORRGIALDQIRKQD 239

DB 241 VVLFQNGSVYRTVYASLETSEGGCPIDPGSSLOKVMYLPPLISNKKORRGIALDQIRKQD 239

QY 301 CLASTTLTAQPDORDAFVGIISYAVKVLFLGALGSELPAELPFVLMHPKPTAKVIHA 360

DB 301 CLASTTLTAQPDORDAFVGIISYAVKVLFLGALGSELPAELPFVLMHPKPTAKVIHA 360

QY 361 DSQADVE 368

DB 361 DSQADVE 368

QY 355 --QPEMET 360

DB 355 --QPEMET 360

RESULT 4

A56607

arrestin homolog - migratory locust

C/Species: *Locusta migratoria* (migratory locust)

C/Date: 11-Aug-1995 #sequence

C/Accession: A56607

R/Ramming, K.; Freitag, J.; Krieger, J.; Breer, H.

Cell. Signal. 5, 69-80, 1993

A/Title: Arrestin-subtypes in insect antennae.

A/Reference number: A56607; MUID:9319955; PMID:8452755

A/Contents: antennae

A/Accession: A56607

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-407 <RAM>

A/Cross-references: UNIPROT:P32122; GB:S57174; NID:9298755; PIDN:AA25860.1; PID:9298756

A/Note: sequence extracted from NCBI backbone (NCBI:127923, NCBI:P127925)

C/Superfamily: arrestin

Query Match

Best Local Similarity 50.4%; Score 993.5; DB 2; Length 407;

Matches 202; Conservative 64; Mismatches 107; Indels 15; Gaps 8;

QY 6 KYEKKCANNGVTLTMGRDPVDHVGVEPIDGIIVLDEYTRDNKRYFGQIVGS 65

DB 6 KYEKKCANNGVTLTMGRDPVDHVGVEPIDGIIVLDEYTRDNKRYFGQIVGS 65

QY 14 KYFKKTTNGKVTVYQKSDFDIDHDHVDPIDGIVVNDNYLR-GRKVFQGLTYYRYGR 72

DB 14 KYFKKTTNGKVTVYQKSDFDIDHDHVDPIDGIVVNDNYLR-GRKVFQGLTYYRYGR 72

QY 66 EEDEVMGAINFOKELCLASBOIYPRPEKSDKEQTKOERLLKKGSAIPTFNISPNAPS 125

DB 66 EEDEVMGAINFOKELCLASBOIYPRPEKSDKEQTKOERLLKKGSAIPTFNISPNAPS 125

QY 73 EEDEVMGAIKTEMTLAKQIYVQ-TKKMKELTPIQERLMKKLGPVAFPTFHPASPS 131

DB 73 EEDEVMGAIKTEMTLAKQIYVQ-TKKMKELTPIQERLMKKLGPVAFPTFHPASPS 131

QY 126 SVTLQOGEEDNDGPGCVSYVYKIFAGESEDTTRHRSSTVTLGIRKIQAPTKOQ-PCT 184

DB 126 SVTLQOGEEDNDGPGCVSYVYKIFAGESEDTTRHRSSTVTLGIRKIQAPTKOQ-PCT 184


```
Db 132 SVTLQPGDDGKPLGVEYVKTWVGDAEKGHRSKRSVTLAIKKLQYAPPTRGRLPSS 191
Qy 185 LVKRDPMLSPELELEVTLDKQLYLHGERIGVNCIRNNSNMKKYKIKAMVQGVVPLF 244
Db 192 LVSGGFPSQKINLEVTLDREIYYHGEKLANVYINNNSKRYKNIKVYVQCEVTMV 251
Qy 245 ONGSRVTVASLETSEGCPIOPGSSLOKVMYLTPLSSNKGRIALDGOIKRQDOCLAS 304
Db 252 -MAQFSNHNVSLETRGECPIFGASFTKVFYLVPCASNKORVGTALDGYIKDDVNLAS 310
Qy 305 TTLAAQ-PDQRDAFVLIISYAVKVLFLGALGSELASLPFVLMHPKPGT----KAKVIH 359
Db 311 STLVSEKNTTDAIGIYISVSLRVKLNCGTLGELQGTDPVFKLHPAPGTERKQAQAI- 369
Qy 360 ADSQADVETFRQD----TIDQASVDPE 383
Db 370 -KKMKSIERTYENSQVAADDDNIVFE 396
```

RESULT 5

```
S11566
arrestin homolog - fruit fly (Drosophila miranda)
C/Species: Drosophila miranda
C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S11566
R:Kriehnan, R.; Ganguly, R.
Nucleic Acids Res. 18, 5894, 1990
A/Title: Nucleotide sequence of the arrestin-like 49 kd protein gene of Drosophila miran
A/Reference number: S11566; MUID:91016944; PMID:2216789
A/Accession: S11566
A/Status: preliminary; translation not shown
A/Molecule type: DNA
A/Residues: 1-401 <KRI>
A/Cross-references: UNIPROT:P19108; EMBL:X54084; NID:97612; PIDN:CAA38019.1; PID:g295738
C/Genetic8:
A/Gene: FlyBase:Dm1r/Arr2
A/Cross-references: FlyBase:FlyBase:F8gn0012552
A/Introns: 286/1; 351/1
C/Superfamily: arrestin
```

```
Query Match 49.5%; Score 977.5; DB 2; Length 401;
Best Local Similarity 48.5%; Pred. No. 3.3e-70;
Matches 190; Conservative 76; Mismatches 113; Indels 13; Gaps 6;
```

```
Qy 1 MVTNFKVFKKCAPNGKVTLYMKGRDPVDHVSQVEPIDGIIVLDEYTRDNKRVFGQIVCS 60
Db 1 MVSIVKVFKKATPNGKVTFLGRDPFDIDLDYCDPVDGVIVVEPEYUK-NRKVFQGLATT 59
Qy 61 FRYGREDEWVGINFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKLGSAIPFTFNIS 120
Db 60 YRGRDEDEWVGVSFSELLISRDEIVPM-TNPMEMTPMQEKLVKRLGSSAHAFPTTFP 118
Qy 121 PNASSTVTLQOEGDNDPCGVSYVVKIFAGESETDTRHRSYTLTGIRKIQFAPTKQO 180
Db 119 PNSSSVTLQOEGDNDKPLGVEYTIKRAFGVDSDDQHKRSWSVLYIKLQVAPLRQO 178
Qy 181 Q-PCTLVKDFMLSPGELIEVTLDKQLYLHGERIGVNCIRNNSNMKKYKIKAMVQGV 239
Db 179 RLPSLSVSKGFTSNGKISLEVTLDREIYYHGEKTAATVQVSNNSKKSIVSIKCFIVQHT 238
Qy 240 DVVLFOGSGYRNTVASLETSEGCPIOPGSSLOKVMYLTPLSSNKGRIALDGOIKRQD 299
Db 239 EITNV-NAQFSKHVAQLETGECPIFGANLTKTFYLIPLAANKKDHGIALDGHLDKD 297
Qy 300 QCLASTTLAQ-PDQRDAFVLIISYAVKVLFLGALGSELASLPFVLMHPKPGT----- 353
Db 298 VNLASSTMVQEGKSTGACGIVISYVRILKNCGTGGEMQTDVPFKLQAPGTIEKR 357
Qy 354 ---KAKVIHADSQADVETFRQDTIDQASVDF 382
Db 358 SNAMKMKKSIQHRNVKGYQDDDDNIVFEDF 389
```

RESULT 6

```
A34856
49k photoreceptor protein - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 09-Jul-2004
C/Accession: A34856
R:Yamada, T.; Takeuchi, Y.; Komori, N.; Kobayashi, H.; Sakai, Y.; Hotte, Y.; Matsumoto, T.
Science 248, 483-485, 1990
A/Title: A 49-kilodalton phosphoprotein in the Drosophila photoreceptor is an arrestin ho
A/Reference number: A34856; MUID:90232360; PMID:2158671
A/Accession: A34856
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-401 <YAM>
A/Cross-references: UNIPROT:P19107; GB:M2141; NID:9158157; PIDN:AAA28833.1; PID:g158158
C/Genetic8:
A/Gene: FlyBase:Arr2
A/Cross-references: FlyBase:F8gn0000121
C/Superfamily: arrestin
```

```
Query Match 49.4%; Score 975.5; DB 2; Length 401;
Best Local Similarity 48.5%; Pred. No. 4.7e-70;
Matches 190; Conservative 74; Mismatches 115; Indels 13; Gaps 6;
```

```
Qy 1 MVTNFKVFKKCAPNGKVTLYMKGRDPVDHVSQVEPIDGIIVLDEYTRDNKRVFGQIVCS 60
Db 1 MVSIVKVFKKATPNGKVTFLGRDPFDIDLDYCDPVDGVIVVEPEYUK-NRKVFQGLATT 59
Qy 61 FRYGREDEWVGINFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKLGSAIPFTFNIS 120
Db 60 YRGRDEDEWVGVSFSELLISRDEIVPM-TNPMEMTPMQEKLVKRLGSSAHAFPTTFP 118
Qy 121 PNASSTVTLQOEGDNDPCGVSYVVKIFAGESETDTRHRSYTLTGIRKIQFAPTKQO 180
Db 119 PNSSSVTLQOEGDNDKPLGVEYTIKRAFGVDSDDQHKRSWSVLYIKLQVAPLRQO 178
Qy 181 Q-PCTLVKDFMLSPGELIEVTLDKQLYLHGERIGVNCIRNNSNMKKYKIKAMVQGV 239
Db 179 RLPSLSVSKGFTSNGKISLEVTLDREIYYHGEKTAATVQVSNNSKKSIVSIKCFIVQHT 238
Qy 240 DVVLFOGSGYRNTVASLETSEGCPIOPGSSLOKVMYLTPLSSNKGRIALDGOIKRQD 299
Db 239 EITNV-NAQFSKHVAQLETGECPIFGANLTKTFYLIPLAANKKDHGIALDGHLDKD 297
Qy 300 QCLASTTLAQ-PDQRDAFVLIISYAVKVLFLGALGSELASLPFVLMHPKPGT----- 353
Db 298 VNLASSTMVQEGKSTGACGIVISYVRILKNCGTGGEMQTDVPFKLQAPGTIEKR 357
Qy 354 ---KAKVIHADSQADVETFRQDTIDQASVDF 382
Db 358 SNAMKMKKSIQHRNVKGYQDDDDNIVFEDF 389
```

RESULT 7

```
B55081
arrestin 2 - bluebottle fly (Calliphora vicina)
C/Species: Calliphora vicina
C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C/Accession: B55081; S44292
R:Planger, A.; Mallick, D.; Whitney, M.; Paulsen, R.
J. Biol. Chem. 269, 26969-26975, 1994
A/Title: Mechanism of arrestin 2 function in rhodometric photoreceptors.
A/Reference number: A55081; MUID:95014564; PMID:7929436
A/Accession: B55081
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-401 <PLA>
A/Cross-references: UNIPROT:P51487; EMBL:X79073; NID:9483583; PIDN:CAA55673.1; PID:g48358
A/Note: The source is designated as Calliphora erythrocephala
C/Superfamily: arrestin
```

```
Query Match 47.8%; Score 942.5; DB 2; Length 401;
```

Best Local Similarity 47.7%; Pred. No. 2e-67;
Matches 187; Conservative 70; Mismatches 122; Indels 13; Gaps 6;

```

OY 1 MYNFKVFKKCAPKPKNTLYNGKRDVFDVHSGVEPIDGIVLDDDEYIRDNKRVGGQVGS 60
DB 1 MYNFKVFKKCAPKPKNTLYNGKRDVFDVHSGVEPIDGIVLDDDEYIRDNKRVGGQVGS 60
OY 61 FRYGEEDEWGNLFOKELCLASEQIYPRPEKSDKEQTKOERLLKLGSGNAIPFTNIS 120
DB 60 YRGRDEDEWGNVFKSKELILCREQIVPM-GNSNMENMTPOEKLVRKLSNNAHPTFHP 118
OY 121 PNASSVTLQOGEEDNDGDCGVSYYVKIFAGESETDRTHRSVTLTGIRKIOFAPTKOQ 180
DB 119 PNPSSVTLQOGEEDNDGDCGVSYYVKIFAGESETDRTHRSVTLTGIRKIOFAPTKOQ 180
OY 181 O-PCTLVKRDPMFLSGELLEVTLDKOLYHGERIGWNCIRNNSNRMVKKIKAMVOGGV 239
DB 179 RUPSSIVSKGFTFSGKISLEVTLDDREIYHGGKAAVQIINNNSKAVNKKIKVFIQHT 238
OY 240 DVLLFQNSYKNTYASLETSEGCPIOPGSSLOKQWYLTPLLSNKKORGIALDQOICL 299
DB 239 EITWV-NAQPSKVAQOLETEGCEPIIRGANLSKTFYLLPLASNNKOBHGIADHKLQED 297
OY 300 QCLASTYLLAQ-PDQDAFVGIISAVYKFLIGALGSELAEPLVLMHPKPTG----- 353
DB 298 VNLASTYVODGKSTGDAQGIYISVRIKLNCTGGEIQIDVFPKLLQAPAPGSVEKR 357
OY 354 --KAVIHADSGADVETFRDPTIDQASVDF 382
DB 358 SNAMKMKSIQHRNTKGYODDDNIVFEDF 389

```

RESULT 8

B34851
beta-arrestin, brain - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 13-Jul-1990 #sequence_revision 21-Oct-1992 #text_change 09-Jul-2004
C/Accession: A34851, J02052
R/Lohse, M.J.; Benovic, J.L.; Codina, J.; Garon, M.G.; Lefkowitz, R.J.
Science 248, 1547-1550, 1990
A/Title: Beta-arrestin, a protein that regulates beta-adrenergic receptor function.
A/Reference number: A34851, MUID:90296080; PMID:2163110
A/Accession: A34851
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-418 <LOH>
A/Cross-references: UNIPROT:P17870; GB:M33601, NID:9162691; PID:AAA30387.1; PID:9162692
A/Experimental source: Brain
R/Source: Mart, R.; Greulich, V.V.; Goldsmith, P.; Bodine, R.C.; Sanders, C.; Donoso, L.A.
J. Biol. Chem. 268, 15640-15648, 1993
A/Title: Polypeptide variants of beta-arrestin and arrestin.
A/Reference number: A47140; MUID:93340166; PMID:8340388
A/Accession: J02052
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-15, 'C', 17-30, 'L', 32, 'K', 34-83, 'Y', 85-176, 'K', 178-206, 'L', 208-333, 342-418
C/Experimental source: lung, spleen
C/Comment: This protein is an arrestin homolog which acts analogously to desensitize the
F/1-418/Products: beta-arrestin, brain #status predicted <MAT>
F/1-333, 342-418/Products: beta-arrestin, lung #status predicted <MAT>

Query Match

Best Local Similarity 45.8%; Score 903.5; DB 2; Length 418;
Matches 179; Conservative 76; Mismatches 112; Indels 13; Gaps 5;

```

OY 6 KVFKKCAPKPKNTLYNGKRDVFDVHSGVEPIDGIVLDDDEYIRDNKRVGGQVGSFYGR 65
DB 7 RVFKKASPNKGLTVYLGKRDVDHIDVDPVGVVLDVEYKKE-RRVYVTLTCAFRIGR 65
OY 66 EDEWGNLFOKELCLASEQIYPRPEKSDKEQTKOERLLKLGSGNAIPFTNISMAAS 125
DB 66 EEDLVLTGLTFKRDVFNANVOSFPAPEDKKPLRLQERLIKKGEGHAYPTTFEIPNLPFC 125

```

OY

```

126 SVTLQOGEEDNDGDCGVSYYVKIFAGESETDRTHRSVTLTGIRKIOFAPTKOQOQPTL 185
DB 126 SVTLQOGEEDNDGDCGVSYYVKIFAGESETDRTHRSVTLTGIRKIOFAPTKOQOQPTL 185
OY 186 VRKQPMLSPELEVTLDKOLYHGERIGWNCIRNNSNRMVKKIKAMVOGGVVFQ 245
DB 186 TTRQFLMSDKPFLHLEASLDKEIYHGERIPISVNAVHTNNKVKYKIKISVQYADICLFN 245
OY 246 NSGYRNTVASLETSEGCPIOPGSSLOKQWYLTPLLSNKKORGIALDQOICL 305
DB 246 TAQKCPVAMEAD--TVAPSTFCVYTLTPFLANNEKRGALDQKLGHEDTNLASS 303
OY 306 TLLAQPDQDAFVGIISAVYKFLF--GALGELSA-----ELPVLMPKPGTK--A 355
DB 304 TLLREGANREILGIVISYKVKLVVSRGGLGLDASSDAVAVELPFLTMHPKKEEPPHR 363
OY 356 KVIHADSGADVETFRDPTID 375
DB 364 EVPEHETPTVNTLIELDTND 383

```

RESULT 9

B3404
beta-arrestin - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: B43404
R/Attamadil, H.; Arriza, J.L.; Aoki, C.; Dawson, T.M.; Codina, J.; Kwatra, M.M.; Snyder
J. Biol. Chem. 267, 17882-17890, 1992
A/Title: Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene family.
A/Accession: B43404
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-418 <AT>
A/Cross-references: UNIPROT:P29066; GB:M91589; NID:6949985; PIDN:AAA74459.1; PID:9203102
C/Suprafamily: arrestin
A/Note: sequence extracted from NCBI backbone (NCBIP:112791)

Query Match

Best Local Similarity 45.8%; Score 903.5; DB 2; Length 418;
Matches 178; Conservative 78; Mismatches 111; Indels 13; Gaps 5;

```

OY 6 KVFKKCAPKPKNTLYNGKRDVFDVHSGVEPIDGIVLDDDEYIRDNKRVGGQVGSFYGR 65
DB 7 RVFKKASPNKGLTVYLGKRDVDHIDVDPVGVVLDVEYKKE-RRVYVTLTCAFRIGR 65
OY 66 EDEWGNLFOKELCLASEQIYPRPEKSDKEQTKOERLLKLGSGNAIPFTNISMAAS 125
DB 66 EEDLVLTGLTFKRDVFNANVOSFPAPEDKKPLRLQERLIKKGEGHAYPTTFEIPNLPFC 125
OY 126 SVTLQOGEEDNDGDCGVSYYVKIFAGESETDRTHRSVTLTGIRKIOFAPTKOQOQPTL 185
DB 126 SVTLQOGEEDNDGDCGVSYYVKIFAGESETDRTHRSVTLTGIRKIOFAPTKOQOQPTL 185
OY 186 VRKQPMLSPELEVTLDKOLYHGERIGWNCIRNNSNRMVKKIKAMVOGGVVFQ 245
DB 186 TTRQFLMSDKPFLHLEASLDKEIYHGERIPISVNAVHTNNKVKYKIKISVQYADICLFN 245
OY 246 NSGYRNTVASLETSEGCPIOPGSSLOKQWYLTPLLSNKKORGIALDQOICL 305
DB 246 TAQKCPVAMEAD--TVAPSTFCVYTLTPFLANNEKRGALDQKLGHEDTNLASS 303
OY 306 TLLAQPDQDAFVGIISAVYKFLF--GALGELSA-----ELPVLMPKPGTK--A 355
DB 304 TLLREGANREILGIVISYKVKLVVSRGGLGLDASSDAVAVELPFLTMHPKKEEPPHR 363
OY 356 KVIHADSGADVETFRDPTID 375
DB 364 EVPEHETPTVNTLIELDTND 383

```

RESULT 10

B46682

beta-arrestin 1, splice form A - human

N/Contains: beta-arrestin 1, splice form B

C/Species: Homo sapiens (man)

C/Date: 22-Sep-1993 #sequence_revision 22-Sep-1993 #text_change 09-Jul-2004

C/Accession: B46682; A46682

R/Parucci, G.; Peracchia, F.; Salliese, M.; Ambrosini, G.; Masini, M.; Rocillo, D.; De Bi

J. Biol. Chem. 268, 9753-9761, 1993

A/Title: Molecular analysis of human beta-arrestin-1: cloning, tissue distribution, and

A/Reference number: A46682; MUID:93252853; PMID:8486659

A/Accession: B46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-418 <PAR1>

A/Cross-references: UNIPROT:P49407; GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347956

A/Note: sequence extracted from NCBI backbone (NCBIP:134014)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Accession: A46682

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Reidues: 1-333, 342-418 <PAR2>

A/Cross-references: GB:L04685; NID:G347956; PIDN:AAA35559.1; PID:G347957

A/Note: sequence extracted from NCBI backbone (NCBIP:134015)

A/Reference number: A47140; MUID:93340166; PMID:8340388

A/Accession: A47140

A/Molecule type: mRNA

A/Reidues: 1-420 <STE>

A/Cross-references: UNIPROT:P32120; GB:L14641

A/Experimental source: Brain

A/Note: sequence extracted from NCBI backbone (NCBIN:136006, NCBIP:136007)

A/Accession: B47140

A/Molecule type: mRNA

A/Reidues: 1-362, 374-420 <ST2>

A/Comment: This protein plays a role in the regulation of G-protein-coupled receptors.

C/Suprafamily: arrestin

F/1-420/Product: arrestin arr3l #status predicted <MAT>

F/1-362,374-420/Product: arrestin arr3s #status predicted <MA2>

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

Query Match 45.0%; Score 887; DB 2; Length 420;

Best Local Similarity 48.3%; Pred. No. 5,86-63;

Matches 167; Conservative 76; Mismatches 97; Indels 4; Gaps 3;

QY 126 SVTLQOGEEDNDGPGCVSYVYKIFAGESEDTDRTRSTVTLGIRKIQAPRTKQGOQPCPTL 185
 Db 127 SVTLQGPEDDTGKAGVDFEIRAFCAKSLSEKSHKNSVRLVIRKQVAPRPEKPPQPSAE 186
 QY 186 VRKDFMLSPGELFVTLTDKQLYHGERIGVNICIRNNSNKKVKKIKAMVQOQGVDFVLFQ 245
 Db 187 TTPHFLMSDRSLHLEASLDKELYYHGERIPVNVHVTNNSSTKVKIKIVSVROYADICLFS 246
 QY 246 NGSYRNTVYASLETSEGCPIOPGSSLQKVMTLPLLSNKKORRGIALDGOIKRDOCLAST 305
 Db 247 TAQKPCPAQLEODD--QVSPSTPCCKYTTTTPPLSDNREKRGALDGLKHEIDTNLASS 304
 QY 306 TLLAQPORDAFGVIIISYAVKYLFLGALGSELSAELPFLTMHPK 351
 Db 305 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349

RESULT 13
 568253
 arrestin isoform 18, erythrocyte - rainbow trout
 C:Species: Oncorhynchus mykiss (rainbow trout)
 C>Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 R:Jahns, R.; Borgese, F.; Lindenthal, S.; Straub, A.; Motaie, R.; Fievet, B.
 Biochem. J. 316, 497-506, 1996
 A:Title: Trout red blood cell arrestin (TRCarr), a novel member of the arrestin family.
 A:Reference number: S68253; MUID:96257743; PMID:8687393
 A:Accession: S68253
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-407 <JAH>
 A:Cross-references: UNIPROT:P51466; GB:U48410; NID:g1215723; PIDN:AA616954.1; PID:g12157
 C:Superfamily: arrestin

Query Match
 Best Local Similarity 44.1%; Score 871; DB 2; Length 407;
 Matches 171; Conservative 82; Mismatches 121; Indels 12; Gaps 5;
 QY 6 KVFKKCAPNGKVTLYMGKDFVDHVSQVEPDIQVLDDEYIRDNKRVGQVCSFRYGR 65
 Db 8 RVKFKSSPNCKTLYVLGKRDPVDHLDVDDVGVLDVDDYLDK--RKVFVTLTCAFYGR 66
 QY 66 EEDVWGINFQKELCLASEQIYRPERKSDKEQTKQERLLKKGNSNAIPFTFNISNAPS 125
 Db 67 EDLDVIGLSFRKDLFITTFQAFPIAEERKANSRLQERLLKKGNSNAIPFTFNISNAPS 125
 QY 126 SVTLQOGEEDNDGPGCVSYVYKIFAGESEDTDRTRSTVTLGIRKIQAPRTKQGOQPCPTL 185
 Db 127 SVTLQGPEDDTGKAGVDFEIRAFCAKSLSEKSHKNSVRLVIRKQVAPRPEKPPQPSAE 186
 QY 186 VRKDFMLSPGELFVTLTDKQLYHGERIGVNICIRNNSNKKVKKIKAMVQOQGVDFVLFQ 245
 Db 187 TTPHFLMSDRSLHLEASLDKELYYHGERIPVNVHVTNNSSTKVKIKIVSVROYADICLFS 246
 QY 246 NGSYRNTVYASLETSEGCPIOPGSSLQKVMTLPLLSNKKORRGIALDGOIKRDOCLAST 305
 Db 247 TAQKPCPAQLEODD--QVSPSTPCCKYTTTTPPLSDNREKRGALDGLKHEIDTNLASS 304
 QY 306 TLLAQPORDAFGVIIISYAVKYLFLGALGSELSAELPFLTMHPK 351
 Db 305 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349
 QY 351 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349
 Db 351 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349
 QY 364 SDPPIIDNLEFETNFSODDDFVE 389

RESULT 14
 A59279
 beta-arrestin 2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 A:Accession: A59279; A43404; I70112

R:Attramadal, H.; Arriza, J.L.; Aoki, C.; Dawson, T.M.; Codina, J.; Kwatra, M.M.; Snyder
 submitted to Genbank, August 1995
 A:Reference number: A59279
 A:Accession: A59279
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-410 <ATT>
 A:Cross-references: UNIPROT:P29067; GB:M91590; NID:g949986; PIDN:AA474460.1; PID:g203104
 A:Experimental source: strain Sprague-Dawley; stage adult; tissue type brain; tissue 11b
 R:Attramadal, H.; Arriza, J.L.; Aoki, C.; Dawson, T.M.; Codina, J.; Kwatra, M.M.; Snyder,
 J. Biol. Chem. 267, 17882-17890, 1992
 A:Title: Beta-arrestin2, a novel member of the arrestin/beta-arrestin gene family.
 A:Reference number: 155423; MUID:94140898; PMID:8308033
 A:Accession: 170112
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 295-410 <CRA>
 A:Cross-references: EMBL:U03627; NID:g458202; PIDN:AA17551.1; PID:g458203
 C:Superfamily: arrestin

Query Match
 Best Local Similarity 48.1%; Score 869.5; DB 2; Length 410;
 Matches 167; Conservative 76; Mismatches 99; Indels 5; Gaps 4;
 QY 6 KVFKKCAPNGKVTLYMGKDFVDHVSQVEPDIQVLDDEYIRDNKRVGQVCSFRYGR 65
 Db 8 RVKFKSSPNCKTLYVLGKRDPVDHLDVDDVGVLDVDDYLDK--RKVFVTLTCAFYGR 66
 QY 66 EEDVWGINFQKELCLASEQIYRPERKSDKEQTKQERLLKKGNSNAIPFTFNISNAPS 125
 Db 67 EDLDVIGLSFRKDLFITTFQAFPIAEERKANSRLQERLLKKGNSNAIPFTFNISNAPS 125
 QY 126 SVTLQOGEEDNDGPGCVSYVYKIFAGESEDTDRTRSTVTLGIRKIQAPRTKQGOQPCPTL 185
 Db 127 SVTLQGPEDDTGKAGVDFEIRAFCAKSLSEKSHKNSVRLVIRKQVAPRPEKPPQPSAE 186
 QY 186 VRKDFMLSPGELFVTLTDKQLYHGERIGVNICIRNNSNKKVKKIKAMVQOQGVDFVLFQ 245
 Db 187 TTPHFLMSDRSLHLEASLDKELYYHGERIPVNVHVTNNSSTKVKIKIVSVROYADICLFS 246
 QY 246 NGSYRNTVYASLETSEGCPIOPGSSLQKVMTLPLLSNKKORRGIALDGOIKRDOCLAST 305
 Db 247 TAQKPCPAQLEODD--QVSPSTPCCKYTTTTPPLSDNREKRGALDGLKHEIDTNLASS 304
 QY 306 TLLAQPORDAFGVIIISYAVKYLFLGALGSELSAELPFLTMHPK 351
 Db 305 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349
 QY 351 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349
 Db 351 TIVEKANKEVIGIIVSYRVKVLV--SRGDVSVLEPFLTMHPK 349

RESULT 15
 T34297
 hypothetical protein F53H8.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 R:Bradshaw, H.; Galling, S.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid F53H8.
 A:Reference number: Z21501
 A:Accession: T34297
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residue: 1-454 <BRA>
A:Cross-references: EMBL:U41023; PIDN:AAA82342.1; CESP:F53H8.2
C:Genetic:
A:Gene: CESP:F53H8.2
A:Introns: 11/2; 61/1; 155/3; 189/2; 237/3; 290/2; 338/2; 389/3; 416/2
C:Superfamily: atreselin

Query Match 43.8%; Score 865; DB 2; Length 454;
Best Local Similarity 46.6%; Pred. No. 3.6e-61;
Matches 174; Conservative 70; Mismatches 101; Indels 28; Gaps 5;

```
Qy      6 KVFKKCAPNGSKVTLVYMGKRDFFVD-----HVSQVEPID-----GIIVLDDEY 46
      11 RVFKKTSFNKSKITTYLGRDPIIDRGDYVDLIGENKGTPEQNVGFILTGLSHGMVLIIDEY 70
Qy      47 IRDNKKVFGQIVCSFRYGREBEDEVMGLNFOKELCLASEQIYPRPEKS-DKEQTKLQERLL 105
      71 IKDNKKVTAHLAFAFRYGREDDLVGLTFRKDLISFTFOVYPQYDKSISRPLSRLOERLK 130
Qy      106 KKLGSNAIPFTFNISPNAPSSVTLQGGEDDNGDPCGVSYVYKIF-----AGESETDRT 156
      131 RKLGAAPFPWFEEVAPKASASVTLQPAAGDTGKPCGVYELKTIYVAVTDGSSGE-KPKKS 189
Qy      159 HRSSTVTLGIRKIQAPPTKOGQCPCTLVKDFMLSPGELBLEVTLDKOLYLHGERIGVNI 218
      190 ALSTVTLAIRKLYAPFESRPQPMVDVSKYFMWSSGLHMEVSLDKEMYHGESISVNV 249
Qy      219 CIRNNSNKKMYKKIYAMVQCGVDVVLFPONGSYRNTVASLETSEGCPIQGSLSQKMYTLTP 278
      250 HIQNSNKKTVKIKIYIIVADICLFTTASYSCEVARIESNEGFPVGGTILSKVPAVCP 309
Qy      279 LLSNKKORRGIALDGOIKRODQCLASTTLLAQPDQDAFGVLIISYAVKVLFLGALGEL 338
      310 LLSNKKDKRGALDGOIKHEDTNLASSITLDSKTSKESLGIVQYRVKVRVAVLGPLNGEL 369
Qy      339 SAELEPVLMPHMKP 351
      370 PAELPFTLTHSKP 382
Db
```

Search completed: February 9, 2005, 23:46:47
Job time : 43 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:23:51 ; Search time 64 Seconds

(without alignments)
3064,474 Million cell updates/sec

Title: US-10-056-405-2

Perfect score: 1973

Sequence: 1 MVNFKVFKKCAPNGKVTLY.....ADVFPRDTRIDQASVDPE 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1973	100.0	383	Q95NF3	O95NF3 anopheles g
2	1973	100.0	417	O7PMG5	O7PMG5 anopheles g
3	1515	76.8	381	O9B1G9	O9B1G9 acaalaphus
4	1499	76.0	381	ARRH_HELV1	P55274 heliothis v
5	1384	70.1	364	ARRA_DROME	P51372 drosophila
6	1335	67.7	363	ARR1_CALVI	P51486 calliphora
7	1254	63.6	245	O6VFP0	O6VFP0 anopheles g
8	1036	52.5	431	O7Q2V9	O7Q2V9 anopheles g
9	1023	51.8	398	O66G06	O66G06 anopheles g
10	1023	51.8	401	O7Q508	O7Q508 anopheles g
11	1002	50.8	470	O9V393	O9V393 drosophila
12	993.5	50.4	407	ARRH_LOCM1	P31122 locusta mig
13	982.5	49.8	401	ARRB_DROME	P319107 drosophila
14	977.5	49.5	401	ARRB_DROME	P319108 drosophila
15	937.5	47.5	400	ARRB_CALVI	P51487 calliphora
16	931	47.2	400	ARRH_LIMPO	P51484 limulus pol
17	931	47.2	419	O86B66	O86B66 halocynthia
18	922.5	46.8	418	O86B66	O86B66 halocynthia
19	919	46.6	418	O64ID8	O64ID8 xenopus lae
20	912.5	46.2	410	O8B7G5	O8B7G5 mus musculu
21	903.5	45.8	418	O8B7G5	O8B7G5 mus musculu
22	903.5	45.8	418	ARR1_BOVIN	P11870 bos taurus
23	903.5	45.8	418	ARR1_RAT	P29066 rattus norv
24	900.5	45.6	418	ARR1_HUMAN	P49407 homo sapien
25	893	45.3	408	O7T2D2	O7T2D2 brachydanio
26	888.5	45.0	435	ARRB_CAEEL	P51485 caenorhabdi
27	887	45.0	420	ARRB_BOVIN	P31120 bos taurus
28	885	44.9	409	ARR2_HUMAN	P31121 homo sapien
29	881	44.7	409	O6ICT3	O6ICT3 homo sapien
30	880.5	44.6	410	ARR1_RABYT	O95223 oryctolagus
31	880	44.6	408	O6DFC4	O6DFC4 xenopus lae

32	878.5	44.5	406	2	O6PFR2	O6PFR2 brachydanio
33	876	44.4	408	2	O6GPR2	O6GPR2 xenopus lae
34	871	44.1	407	1	ARR1_ONCMY	P51466 oncotrynychu
35	869.5	44.1	410	1	ARR2_MOUSE	O91Y14 mus musculu
36	869.5	44.1	410	1	ARR2_MOUSE	P29067 rattus norv
37	863	43.7	415	1	ARR3_ONCMY	P51468 oncotrynychu
38	861	43.6	405	1	ARR2_ONCMY	O9PTE7 ambystoma t
39	853	43.2	392	2	O9PTE7	O9PTE7 ambystoma t
40	810	41.1	387	1	ARRC_XENILA	P51483 xenopus lae
41	791	40.1	392	1	O6NUV2	O6NUV2 brachydanio
42	785.5	39.8	389	1	ARRC_RANCA	P51481 rana catesb
43	784.5	39.8	389	1	ARRC_RANPI	P51482 rana pipien
44	783.5	39.7	403	1	ARRS_MOUSE	P20443 mus musculu
45	781.5	39.6	403	1	ARRS_RAT	P15887 rattus norv

ALIGNMENTS

RESULT 1					
ID	Q95NF3	PRELIMINARY;	PRT;	383 AA.	
AC	Q95NF3				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			
DT	05-JUL-2004	(TREMBLrel. 27, Last annotation update)			
DE	Arrestin.				
GN	Name=Arri.				
OS	Anopheles gambiae (African malaria mosquito).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.				
OX	NCBI_Taxid=7165;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Antennae, and Head;				
RC	MEDLINE=21819361; PubMed=11792843; DOI=10.1073/pnas.02505499;				
RX	MEDLINE=21819361; PubMed=11822731;				
RA	Ricci I., Santolamazza F., Costantini C., Pavia G.;				
RA	Carlson J.R., Zwiabed L.U.;				
RT	"Visual arrestin in olfactory pathways of Drosophila and the malaria				
RT	vector mosquito Anopheles gambiae."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:1633-1638(2002).				
DR	EMBL; AJ304409; CAC39103.2;				
DR	EMBL; AY017417; AAC54081.1;				
DR	HSSP; P17870; 1G4M.				
DR	GO; GO:0007600; P:sensory perception; IEA.				
DR	GO; GO:0007165; P:signal transduction; IEA.				
DR	InterPro; IPR000698; Arrestin.				
DR	InterPro; IPR011022; Arrestin_C.				
DR	InterPro; IPR011021; Arrestin_N.				
DR	Pfam; PF02752; Arrestin_C; 1.				
DR	Pfam; PF00339; Arrestin_N; 1.				
DR	PRINTS; PRO0309; ARRESTIN.				
DR	PROSITE; PS00295; ARRESTIN; 1.				
SQ	SEQUENCE 383 AA; 42809 MW; FA946438592B53E CRC64;				
Query Match 100.0%; Score 1973; DB 2; Length 383;					
Best local similarity 100.0%; Pred. No. 1.1e-144;					
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MVNFKVFKKCAPNGKVTLYGKRDVDVSGVPEIDGIWLDDEYTRDNKKVFGQIVCS	60		
DB	1	MVNFKVFKKCAPNGKVTLYGKRDVDVSGVPEIDGIWLDDEYTRDNKKVFGQIVCS	60		
QY	61	FRYGRDEWVGILNFKELCLASFOIYPRPEKSDKEQTKLQERLLKLGSGNAIPFTNIS	120		

```

Db      61 FRYGREEDVWGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKKGSNALPFTFNIS 120
Qy      121 FNAPSSVTLQOGEEDNDGDPGVSYVYKIFAGESETDRTHRSSTVTLGIRKIOFAPTKOQ 180
Db      121 FNAPSSVTLQOGEEDNDGDPGVSYVYKIFAGESETDRTHRSSTVTLGIRKIOFAPTKOQ 180
Qy      181 QPCTVLRDPMLSPELELEVTLDKOLYHGERIGVNICIRNNSNKVKKIKAMVOQGV 240
Db      181 QPCTVLRDPMLSPELELEVTLDKOLYHGERIGVNICIRNNSNKVKKIKAMVOQGV 240
Qy      241 VLFONGSYRNTVAASLETSEGCPIQPGSSLOKWYLTPLSSNKORRGIALDGOIKRODQ 300
Db      241 VLFONGSYRNTVAASLETSEGCPIQPGSSLOKWYLTPLSSNKORRGIALDGOIKRODQ 300
Qy      301 CLASTTLLAOPDORBAFVITISYAVKVLFLGALGSELAEPLVLMHPKPGTKAKYIHA 360
Db      301 CLASTTLLAOPDORBAFVITISYAVKVLFLGALGSELAEPLVLMHPKPGTKAKYIHA 360
Qy      361 DSQADVETFRDPTIDQOASVDPE 383
Db      361 DSQADVETFRDPTIDQOASVDPE 383

```

RESULT 2

```

ID      07PMG5      PRELIMINARY;      PRT;      417 AA.
AC      07PMG5;
DT      01-MAR-2004 (TEMBLrel. 26, Created)
DT      01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DE      ENSANGP0000012569 (Fragment)
GN      Name=ENSANGP0000012569 (Fragment)
OS      Anopheles gambiae str. PEST.
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OX      Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
RN      NCBI_Taxid=180454;
RP      (1)
RC      SEQUENCE FROM N.A.
RA      Anopheles gambiae.
RL      Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC      -! CAUTION: The sequence shown here is derived from an
CC      EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
DR      HSBP; A18801008980; EMBL3874.2; -.
DR      GO; GO:0007600; P:sensory perception; IEA.
DR      GO; GO:0007165; P:signal transduction; IEA.
DR      InterPro; IPR000698; Arrestin.
DR      InterPro; IPR011021; Arrestin_C.
DR      Pfam; PF00339; Arrestin_N; 1.
DR      PRINTS; PR00309; ARRESTIN.
DR      PRODOM; PD002098; ARRESTIN.
DR      PROSITE; PS00295; ARRESTINS; 1.
FT      NON TER
SQ      SEQUENCE 417 AA; 46591 MW; 4EB9BA3C32257FC0 CRC64;

```

Query Match Best Local Similarity 100.0%; Score 1973; DB 2; Length 417; Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MYNFKYFKKCAPNGKVTLYMGKRDVFDHVGVEPIDGIVLVDDEYIRDNKRVGQIVCS 60
Db      35 MYNFKYFKKCAPNGKVTLYMGKRDVFDHVGVEPIDGIVLVDDEYIRDNKRVGQIVCS 60
Qy      61 FRYGREEDVWGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKKGSNALPFTFNIS 120
Db      95 FRYGREEDVWGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKKGSNALPFTFNIS 120
Qy      121 FNAPSSVTLQOGEEDNDGDPGVSYVYKIFAGESETDRTHRSSTVTLGIRKIOFAPTKOQ 180

```

```

Db      155 FNAPSSVTLQOGEEDNDGDPGVSYVYKIFAGESETDRTHRSSTVTLGIRKIOFAPTKOQ 214
Qy      181 QPCTVLRDPMLSPELELEVTLDKOLYHGERIGVNICIRNNSNKVKKIKAMVOQGV 240
Db      215 QPCTVLRDPMLSPELELEVTLDKOLYHGERIGVNICIRNNSNKVKKIKAMVOQGV 240
Qy      241 VLFONGSYRNTVAASLETSEGCPIQPGSSLOKWYLTPLSSNKORRGIALDGOIKRODQ 300
Db      241 VLFONGSYRNTVAASLETSEGCPIQPGSSLOKWYLTPLSSNKORRGIALDGOIKRODQ 300
Qy      301 CLASTTLLAOPDORBAFVITISYAVKVLFLGALGSELAEPLVLMHPKPGTKAKYIHA 360
Db      301 CLASTTLLAOPDORBAFVITISYAVKVLFLGALGSELAEPLVLMHPKPGTKAKYIHA 360
Qy      361 DSQADVETFRDPTIDQOASVDPE 383
Db      361 DSQADVETFRDPTIDQOASVDPE 383

```

RESULT 3

```

ID      09BIG9      PRELIMINARY;      PRT;      381 AA.
AC      09BIG9;
DT      01-JUN-2001 (TEMBLrel. 17, Created)
DT      01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE      Arrestin.
GN      Name=ar1;
OS      Ascalaphus macaronius.
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OX      Neoptera; Endopterygota; Neuroptera; Ascalaphidae; Ascalaphus.
RN      NCBI_Taxid=16496;
RP      (1)
RC      SEQUENCE FROM N.A.
RA      MEDLINE=21164354; PubMed=1128706; DOI=10.1016/S0014-5793(01)02287-6;
RA      "UV-light-dependent binding of a visual arrestin 1 isoform to
RT      photoreceptor membranes in a neuropteran (Ascalaphus) compound eye.";
RL      FBS Lett. 493:112-116(2001).
DR      EMBL; AJ303080; CAC36938.1; -.
DR      HSBP; P17870; IGFM.
DR      GO; GO:0007600; P:sensory perception; IEA.
DR      GO; GO:0007165; P:signal transduction; IEA.
DR      InterPro; IPR000698; Arrestin.
DR      InterPro; IPR011021; Arrestin_C.
DR      Pfam; PF02752; Arrestin_N; 1.
DR      PRINTS; PR00309; ARRESTIN.
DR      PROSITE; PS00295; ARRESTINS; 1.
SQ      SEQUENCE 381 AA; 42789 MW; 451DEF465969857D CRC64;

```

Query Match Best Local Similarity 76.8%; Score 1515; DB 2; Length 381; Matches 289; Conservative 46; Mismatches 41; Indels 6; Gaps 4;

```

Qy      1 MYNFKYFKKCAPNGKVTLYMGKRDVFDHVGVEPIDGIVLVDDEYIRDNKRVGQIVCS 59
Db      1 MYNFKYFKKCAPNGKVTLYMGKRDVFDHVGVEPIDGIVLVDDEYIRDNKRVGQIVCS 59
Qy      60 FRYGREEDVWGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKKGSNALPFTFNIS 119
Db      61 FRYGREEDVWGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKKGSNALPFTFNIS 119
Qy      120 SPNAPSSVTLQOGEEDNDGDPGVSYVYKIFAGESETDRTHRSSTVTLGIRKIOFAPTKOQ 178
Db      119 SPNAPSSVTLQOGEEDNDGDPGVSYVYKIFAGESETDRTHRSSTVTLGIRKIOFAPTKOQ 178
Qy      178 QPCTVLRDPMLSPELELEVTLDKOLYHGERIGVNICIRNNSNKVKKIKAMVOQGV 239
Db      179 QPCTVLRDPMLSPELELEVTLDKOLYHGERIGVNICIRNNSNKVKKIKAMVOQGV 239

```



```
Qy 240 DVVLFONGSYRNTVALETSBGCPIOPGSSLOKVMYLLPILLSNKKORGIALDGOIRKOD 299
    |||||
Db 239 DVVLFONGQORTTIAHLETFQEGCPIOPGSAQKMYLLPILLSNKKDRGIALDGOIRKOD 298
Qy 300 QCLASTLLAOPDRDAFGVIISYAVKVLFLGALGSELSALPFLVLMHPKGTAKYVIH 359
    |||||
Db 299 TNLASTLLANPEQRDAFGIIVSYAVKVLGALGELALPFLIMHPRP-ERNKFIH 357
Qy 360 ADSQADVETFRDPTIDQASVD 381
    |||||
Db 358 ADSQADVEMFRQDTVD--PSVD 377

RESULT 4
ARRH_HELVI STANDARD; PRT; 381 AA.
ID _ARRH_HELVI
AC P55274;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Arreclin homol.9.
OS Heliothis virescens (Noctuid moth) (Owllet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditypsia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
NCBI_TaxId=7102;
OX
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Antenna;
RX MEDLINE=3139955; PubMed=8452755; DOI=10.1016/0898-6568(93)90009-B;
RA Raming K., Freitag J., Krieger J., Bree H.;
RT "Arreclin-subtypes in insect antennae.";
RL Cell. Signal. 5:69-80(1993).
CC -1- SIMILARITY: Belongs to the arreclin family.
DR HSP; B56607; B56607.
DR HSSP; P17870; IG4M.
DR InterPro: IPR000698; Arreclin.
DR InterPro: IPR011022; Arreclin.C.
DR InterPro: IPR011021; Arreclin.N.
DR Pfam; PF02732; Arreclin_C_1.
DR Pfam; PF00339; Arreclin_N_1.
DR PRINTS; PR00309; ARRESTIN.
DR Prodom; PD002099; Arreclin_2.
DR PROSITE; PS00295; ARRESTINS; 1.
KW Sensory transduction.
SQ SENSOURCE 381 AA; 42747 MW; 84BB92B1BB3DA573 CRC64;

Query Match 76.0%; Score 1499; DB 1; Length 381;
Best Local Similarity 72.7%; Pred. No. 6,5e-108;
Matches 277; Conservative 51; Mismatches 51; Indels 2; Gaps 2;
```

```
Db 300 TLASTLLLPDQDRDAFGIIVSYAKVLYLGISELVALPFLIMHPKRG-RVKMIIHA 358
    |||||
Qy 361 DSQADVETFRDPTIDQASVD 381
    |||||
Db 359 DSQADVEMFRDPTVHQSVE 379

RESULT 5
ARRA_DROME STANDARD; PRT; 364 AA.
ID _ARRA_DROME
AC P15372; Q9VUA8;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Phorestin II (Arrestin A) (Arrestin 1)
DE Phorestin II (Arrestin A) (Arrestin 1)
GN Name=Arri; Synonyms=Artra; ORFName=CG5711;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
OX
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=90138926; PubMed=2105491;
RA Hyde D.R., Mecklenburg K.L., Pollock J.A., Vintelic T.S., Benzer S.;
RT "Twenty Drosophila visual system cDNA clones: one is a homolog of
human arrestin.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1008-1012(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherter S.E., Li P.W., Hoekins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazer J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfennoch C., Baldwin D.,
RA Balieu R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintakov S.,
RA Borkova D., Botchan M.R., Bouck J., Broksstein P., Brotlier P.,
RA Burtie K.C., Buesam D.A., Butler H., Cadieu E., Cantler A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Paloe B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foele C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Goelzer A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mleishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy U., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
```

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN [4]
 RP GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tsip J.L., Whitfield J., Bayraktaroglu L., Bertan B.P.,
 RA Bellencontre B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Stapleton M., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Lewis S.E., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RL systematic review,"
 RN Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley; TISSUE=Head;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Garlin H., Krommiller B., Pacleab J.M., Park S., Wan K.H.,
 RT Rubin G.M., Celinker S.E.,
 RL "A *Drosophila* full-length cDNA resource,"
 RN Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 RP PHOSPHORYLATION.
 RX MEDLINE=91282780; PubMed=1905538;
 RA Matsuno H., Yamada T.,
 RT "Phosphatases I and II: arrestin homologs which undergo differential
 RT light-induced phosphorylation in the *Drosophila* photoreceptor in
 RL vivo,"
 RN Biochem. Biophys. Res. Commun. 177:1306-1312(1991).
 RP FUNCTION.
 RX MEDLINE=93303590; PubMed=8316831;
 RA Dolph P.J., Rangamathan R., Colley N.J., Hardy R.W., Socolich M.,
 RT Zuker C.S.,
 RT "Arrestin function in inactivation of G protein-coupled receptor
 RL rhodopsin in vivo,"
 RN Science 260:1910-1916(1993).
 CC -1- FUNCTION: Regulates photoreceptor cell deactivation. Arr1 and Arr2
 CC for the termination of the phototransduction cascade.
 CC -1- TISSUE SPECIFICITY: Expressed specifically and abundantly in the
 CC photoreceptors. Inner and outer segments, and the inner plexiform
 CC -1- PTM: Phosphorylated, but does not undergo light-induced
 CC phosphorylation.
 CC -1- SIMILARITY: Belongs to the arrestin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 DR EMBL: M30177; NOT ANNOTATED CDS.
 DR EMBL: M30140; AAA28350.1;
 DR EMBL: AE003657; AARS344.1;
 DR EMBL: AY061824; AAL27635.1;
 DR PIR: A34867; A34867.
 DR HSSP: P17870; IGAM.
 DR Interact: P15372;
 DR FLYBase: FBgn0000120; Arr1.
 DR GO: GO:0005624; C:membrane fraction; IDA.
 DR GO: GO:0016028; C:rhododerm; IDA.
 DR GO: GO:0016060; P:metarhodopsin inactivation; IGI.
 DR InterPro: IPR000698; Arrestin.
 DR InterPro: IPR011021; Arrestin_C.

DR InterPro: IPR011021; Arrestin_N.
 DR Pfam: PF02752; Arrestin_C; 1.
 DR Pfam: PF00339; Arrestin_N; 1.
 DR PRINTS: PR00309; ARRESTIN.
 DR PRODOM: PD002099; ARRESTIN.
 DR PROSITE: PS00295; ARRESTIN; 2.
 KW Phosphorylation; Sensory transduction; Vision.
 SQ SEQUENCE 364 AA; 4077 MW; 0DC764CF890FC2 CRC64;
 Query Match 70.1%; Score 1384; DB 1; Length 364;
 Best Local Similarity 70.3%; Pred. No. 5.1e-99;
 Matches 258; Conservative 51; Mismatches 54; Indels 4; Gaps 2;
 QY 1 MYNFVFVKKAPNGKVTLYMGKRDVVDHSGVEPIDGIIVLDDERYLRNRYFGQVCS 60
 DB 1 MYNFVKFKK...
 QY 61 PRYGESEDEWNGINPQKEI...
 DB 61 PRYGESEDEWNGINPQKEI...
 QY 121 PNASVTLTQCGEDNDGPPCGVSYVYKIPAGESETDRTRRSVTTLGTRKIQFAPTKQG 119
 DB 120 PSSPASVVLQKASDESQPCGVYFVKIFGDSDCDRSHRSTINLGIKQVAFPTKGI 119
 QY 181 QPCTLVKDFMISRGEL...
 DB 180 QPCTLVKDFMISRGEL...
 QY 241 VVLFONGSYRTVYASLETSEGCPIOGSSLOKWTLPPLSSKORRGALNGQIKRQDQ 239
 DB 240 VVLFONGSYRTVYASLETSEGCPIOGSSLOKWTLPPLSSKORRGALNGQIKRQDQ 239
 QY 301 CLASTTLLAQPDQDAFGVITSYAVKTKFLGALGSELAEIPFLMPKPGCTAKYTHA 300
 DB 300 ALASTTLLAQPDQDAFGVITSYAVKTKFLGALGSELAEIPFLMPKPGCTAKYTHA 300
 QY 361 DSGADVE 367
 DB 357 EAEGSIE 363
 RESULT 6
 ID ARR1_CALVI STANDARD; PRT; 363 AA.
 AC P51486;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 GN Name=ARR1;
 OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
 OC Neoptera; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Eukaryota; Enderlepterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 OC Calliphoridae; Calliphora.
 OX NCBI_TaxID=7373;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina.
 RX MEDLINE=95014564; PubMed=7929436;
 RA Planger A., Mallick D., Whitney M., Paulsen R.,
 RT "Mechanism of arrestin 2 function in rhododermic photoreceptors,"
 RL J. Biol. Chem. 269:26969-26975(1994).
 CC -1- SIMILARITY: Belongs to the arrestin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).

DR InterPro: IPR011021; Arrestin_N.
DR Pfam: PF02752; Arrestin_C; 1.
DR PRINTS: PR00309; ARRESTIN.
DR PRODOM: PD002099; ARRESTIN.
DR PROSITE: PS00295; ARRESTIN; 2.
FT NON TER 1 431 1
SQ SEQUENCE 431 AA; 47810 MW; 85C110785A815916 CRC64;
Query Match
Best Local Similarity 52.8%; Score 1036; DB 2; Length 431;
Matches 191; Conservative 75; Mismatches 85; Indels 2; Gaps 2;
QY 6 KVEKCAKNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDBYIRDRNRKVGQIVCSFPRGR 65
DB 7 RVKSSSSNGKITVYLGKRDVDFVHSGVEPIDGIVLDDBYIRDRNRKVGQIVCSFPRGR 65
QY 66 EEDVWGNFQKEICLASEQIYPRPKSDKEQTKLOERLLKKGASNAIPFTFNIS 125
DB 66 EEDVWGNFQKEICLASEQIYPRPKSDKEQTKLOERLLKKGASNAIPFTFNIS 125
QY 66 EEDVWGNFQKEICLASEQIYPRPKSDKEQTKLOERLLKKGASNAIPFTFNIS 125
DB 66 EEDVWGNFQKEICLASEQIYPRPKSDKEQTKLOERLLKKGASNAIPFTFNIS 125
QY 126 SVTLQGGEDNDGPPCGVSYVYKIFAGESETDTRHRTSTVTLGIRKIQAPFTQGGQPC 124
DB 125 SVTLQGGEDNDGPPCGVSYVYKIFAGESETDTRHRTSTVTLGIRKIQAPFTQGGQPC 124
QY 186 VRDPMFLSPGELELVTLDKOYLHGERIGVNICIRNNSNMVKKIKAMVQGVVYLFQ 185
DB 185 VRDPMFLSPGELELVTLDKOYLHGERIGVNICIRNNSNMVKKIKAMVQGVVYLFQ 185
QY 246 NSGYRTVYSLSTSGCPIQGGSLQKVMYLPPLSSNKGRRGIALDQIRROQCLAST 244
DB 245 NSGYRTVYSLSTSGCPIQGGSLQKVMYLPPLSSNKGRRGIALDQIRROQCLAST 244
QY 306 TLLADPDRQAFGVITSYAVKVLFLGALGSLAEPLFVLMHPKPGTAKY 305
DB 305 TLLADPDRQAFGVITSYAVKVLFLGALGSLAEPLFVLMHPKPGTAKY 305
QY 305 TLLADPDRQAFGVITSYAVKVLFLGALGSLAEPLFVLMHPKPGTAKY 305
DB 305 TLLADPDRQAFGVITSYAVKVLFLGALGSLAEPLFVLMHPKPGTAKY 305
RESULT 9
Q66GUG PRELIMINARY; PRT; 398 AA.
AC 066GUG; PRT; 398 AA.
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DE 25-OCT-2004 (Tremblrel. 28, Last sequence update)
GN Name=ARR2; Anopheles 2-like protein Arr2.
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PEST;
RA MEDLINE=23159591; PubMed=14986925;
RA Merrill C.E.; Pites R.J.; Zweibel L.J.;
RT "Molecular characterization of arrestin family members in the malaria
vector mosquito, Anopheles gambiae.";
RL Insect Mol. Biol. 12:641-650(2003).
DR EMBL: BK000996; DA000888.1;
DR InterPro: IPR000698; Arrestin.
DR InterPro: IPR011021; Arrestin_C.
DR Pfam: PF02752; Arrestin_C; 1.
DR PRINTS: PR00309; ARRESTIN_N.
DR PRODOM: PD002099; ARRESTIN.
DR PROSITE: PS00295; ARRESTIN; 2.
FT NON TER 1 431 1
SQ SEQUENCE 398 AA; 44536 MW; 73DD73FB34F01418 CRC64;
Query Match
Best Local Similarity 52.8%; Score 1023; DB 2; Length 398;
Matches 191; Conservative 75; Mismatches 85; Indels 2; Gaps 2;

Matches 204; Conservative 65; Mismatches 111; Indels 10; Gaps 7;
QY 1 MYNFKYKCAKNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDBYIRDRNRKVGQIVCS 60
DB 1 MYNFKYKCAKNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDBYIRDRNRKVGQIVCS 60
QY 61 FRYGDEDEWGNFQKEICLASEQIYPRPKSDKEQTKLOERLLKKGASNAIPFTFNIS 120
DB 60 FRYGDEDEWGNFQKEICLASEQIYPRPKSDKEQTKLOERLLKKGASNAIPFTFNIS 120
QY 121 PNASSVTLQGGEDNDGPPCGVSYVYKIFAGESETDTRHRTSTVTLGIRKIQAPFTQGG 118
DB 119 PNASSVTLQGGEDNDGPPCGVSYVYKIFAGESETDTRHRTSTVTLGIRKIQAPFTQGG 118
QY 181 Q-PCITLVKMDLSPGELELVTLDKOYLHGERIGVNICIRNNSNMVKKIKAMVQGV 178
DB 179 RLPSSLVSKGFTTSQKINLEVTLDREIYHGEKIANIYVNNRSKTKYSIKCFVYOH 238
QY 240 DVLPFGSYRNTVASLSTSGCPIQGGSLQKVMYLPPLSSNKGRRGIALDQIRROQCLAST 244
DB 239 DVLPFGSYRNTVASLSTSGCPIQGGSLQKVMYLPPLSSNKGRRGIALDQIRROQCLAST 244
QY 300 QCLASTTLLADPDRQAFGVITSYAVKVLFLGALGSLAEPLFVLMHPKPGTAKY 305
DB 298 QCLASTTLLADPDRQAFGVITSYAVKVLFLGALGSLAEPLFVLMHPKPGTAKY 305
QY 357 VIALKQMSIRKHYNSHYADDDNIVFE 387
DB 358 VIALKQMSIRKHYNSHYADDDNIVFE 387
RESULT 10
Q70508 PRELIMINARY; PRT; 401 AA.
AC 070508; PRT; 401 AA.
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DE AGP5798 (Fragment).
GN Name=agc52387; ORFNames=ENSG0000018139;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
CC Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC Preliminary data.
DR EMBL: AA01008960; EAA10755.1; -.
DR HSSP: P17870; 1G4M.
DR GO: GO:0007600; P:genery perception; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.
DR InterPro: IPR000698; Arrestin.
DR InterPro: IPR011021; Arrestin_C.
DR Pfam: PF02752; Arrestin_C; 1.
DR PRINTS: PR00309; ARRESTIN_N.
DR PRODOM: PD002099; ARRESTIN.
DR PROSITE: PS00295; ARRESTIN; 2.
FT NON TER 1 401 1
SQ SEQUENCE 401 AA; 44862 MW; F307885989A4C4E8 CRC64;
Query Match
Best Local Similarity 52.8%; Score 1023; DB 2; Length 401;
Matches 204; Conservative 65; Mismatches 111; Indels 10; Gaps 7;
QY 1 MYNFKYKCAKNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDBYIRDRNRKVGQIVCS 60
DB 1 MYNFKYKCAKNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDBYIRDRNRKVGQIVCS 60

Db 4 MVVAVKFKSA PNGKLT VYLGRDPLDHTDYPDIDGIVLDEBYLR-GRKVEGOLITT 62
 Qy 61 FRVREDEWVWGLNFOKEILCLASBOIYRPREKSKXEOQKQOERLLKLUKLSNAIFFTNIS 120
 Db 63 YRVREDEWVWGVFSKEMVLTKEOILYPM-BNANMEMTWMOERLVKLGANAAPPTPHP 121
 Qy 121 PNASVYTLQOGEEDNCPGCVSYVVKI FAGESETDRTHRSSTYTLGIKRIQFAPTRQOQ 180
 Db 122 SMAPSSVTLQAGEDDTKPLGVEYAIKAVHGEDSDSKGKHSATYLTITKLUQYAPVBSGR 181
 Qy 181 Q-PCTLYRQDPMLSPELELEVTLDKOLYLHGERIGVNICIRANSNKVKIKAMVOGV 239
 Db 182 RLPSLSLXSKGFTFGQKINLEVTLDREIYHGKELIANIIVTNNSRKTVKSIKCFVVOHC 241
 Qy 240 DVYLFQNGSVYNTASLETSGCPIDGSSLOKMMVYLTPLLSNKKORGLLDGQIRPD 299
 Db 242 EVTVV-NAQFSKHLIASLRETRGCPITPEASTKSPFLPLSSNKKDRGIALDGHLEDD 300
 Qy 300 QCLASTTLLAQPD-QRDAFVGIISYAVVKLFLGALGELSAPLPLVLMHPKPT--KAK 356
 Db 301 VNLAASSTLISGCKPSDAMGIVISYSLRVKINCSTLGGELQTDVPLKLMNAPGSEVERER 360
 Qy 357 VHAHQADVETFRQDT--IDQASVDFE 383
 Db 361 VNALKKMKSIERHRYENSHYADDDNIVFE 390
 RESULT 11
 Q9V393 ID Q9V393 PRELIMINARY; PRT; 470 AA.
 AC Q9V393;
 DT 01-MAY-2000 (TREMblrel. 13. Created)
 DT 01-MAY-2000 (TREMblrel. 13. Last sequence update)
 DT 25-OCT-2004 (TREMblrel. 28. Last annotation update)
 DE CG1487-PA (Kurtz arrestin) (LD31082p).
 GN Name=Krtz; ORFNames=CG1487;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.B.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abtill J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunham B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
 RA Fogle C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hootin D., Houson K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jatala M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lakso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamoe I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodrager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yen R.F., Zaveri U.S., Zhan M., Zhang X., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2242605; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Patel J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskaas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RL melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RL a genomic perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426059; PubMed=12537572;
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochman S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RX Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RX Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RL [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20341329; PubMed=10880488;
 RA Roman G., He J., Davis R.L.;
 RT "Kurtz, a novel nonvisual arrestin, is an essential neural gene in
 RL *Drosophila*.";
 RL Genetics 155:1281-1295(2000).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacle J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A0003780; AAF57212.1; -;
 DR EMBL: AF221066; AAF32365.1; -;
 DR EMBL: AY118946; AAM50806.1; -;
 DR HSSP: P17870; 1G4M.
 DR IntAct: Q9V393; -.

Query Match	50.8%;	Score 1002;	DB 2;	Length 470;
Best Local Similarity	53.7%;	Pred. No. 3.2e-69;		
Matches 187;	Conservative 70.00%;			

QY 6 KVKKCKAPNGKVTLLYMGKRDPEVDHVGSEPIOSIVLDEDEYTRDNRAKVGQIYCSRYR 4; Gaps 3
Db 48 RFFKSSSSNGKITYVLGKRDPEVDHVTVDHIDGVVEIPDEYKD-RVYFQVLAARIGR 1
QY 66 EDEYWGAINFOEELCLASBQIYRPRPKSDKEOTKIQEELKKGASNAIPPTENISPNAS 106
Db 107 EDLDVLGILTRKDIYLHNEGIYR-PWQDRPMWTRLQEELIKKGASNAIPFEYVPPCYA 125
QY 126 SVTLQGGEDDNGRPGCVSYVVKIPAGESETRTRRSTVTLGIRIQAPATPOGQOPOTL 165
Db 166 SVSLQAPRPGVNGSCGVDELKAPFGEVNEDEKPRKNSVTLTRKMYAPSKVGEOSIE 185
QY 186 VRKDPMLSPPELETLVTLDDKLYLHGERIYGINICIRNNSKPKVKIKAMVOOGVYDVLFPQ 225
Db 226 VSSEKFMWKPKKIHLETLDDKELYLHNGEKIYSNVHANNSNRKIVYKIKVCARQFADICTFS 245
QY 246 NSGYRNTVYASLTSSECPRIQSSLSQVMTLPTLLSSNKQRRGIALDSQILRODCCLAST 285
Db 286 TAOYKSVVAELISESDGCVPARGPFLSLVPELCPLLANKKDWGALDQGHEDYTLNAS 305
QY 306 TLLAOPDRDAFGVYISAVKMYKLTGA-LGGELSAELPVLNHPRP 351
Db 346 TLTITPAGRESLGIMYHKVKVTLTLLSSPLINGDLVVELPPTLLNHPRP 393

RESULT 12	
ARRH_LOCM1	
ID ARRH_LOCM1	STANDARD;
AC P32122;	PRT; 407 AA

BT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
ID 15-JUL-1999 (Rel. 38, Last annotation update)
DE Arrestin homolog
LO Locusta migratoria (Migratory locust).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Acroptera; Orthopteroidea; Orthoptera; Cellifera; Acridomorpha;
NC Acrididae; Acrididae; Oedipodinae; Locusta.
NX NCBI_TaxId=7004;
[1]
SEQUENCE FROM N.A.
CC TISSUE=Antenna;
CC MEDLINE=93199955; PubMed=8452755; DOI=10.1016/0898-6568(93)90009-B;
CC Raming K., Freitag J., Krüger J., Brer H.,
CC "Arrestin-subtypes in insect antennae.",
CC Cell. Signal. 5:59-60(1993).
-I- SIMILARITY

CC ----- belongs to the arrestin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR
DR EMBL; S57174; AAB25860.1; -
DR PIR; A56607; A56607

```

DR  HSSP; P17870; 1G4M.
DR  InterPro; IPR000698; Arrestin.
DR  InterPro; IPR011022; Arrestin_C.
DR  InterPro; IPR011021; Arrestin_N.
DR  Pfam; PF02752; Arrestin_C_1.
DR  Pfam; PF00339; Arrestin_N_1.
DR  PRINTS; PR00309; ARRESTIN.
DR  ProDom; PD002099; Arrestin.
DR  PROSITE; PS00295; ARRESTINS; 1.
KW  Sensor transduction.
CC  SEQUENCE

```

Query Match	50.4%	Score 993.5;	DB 1;	Length 407;
Best Local Similarity	52.1%;	Pred. No. 1.2e-68;		
Matches 202;	Conservative			

Qy	6	KYFKCKAPNGKYTL YMGKRDVPHVSGVEPIDIGIVLDDDEYTRNNRNFQOICSPRYER	65
Db	14	KYFKKTTTNGKAVYVYLKGRDFIDHLDHVDIDIGIVVNDYLR GRVYFQSLTTRYGR	72
Qy	66	EEDVYMGINFQKELCLASBQIYVPRPKDKEQTKLGERLLKUGSNAIPFTFNISPMPS	125
Db	73	EEDVYMGKFTKEMVYLAKQIWPQ TKKMKELTPIORLKKMKKGPAFFTHFPASSPS	131
Qy	126	SVTLQOQEDDNGPQGVSVYVYKIPGSESDRTHTRSYTLVIGRIKOPATKQGOO PCT	184
Db	132	SVTLQPCDDQKGPAGVEYSVKTAVGWHAEKKGKRSVATLTKLQTPAPPRRRPSS	191
Qy	185	LYRKDPMLSRGLLEVTLDKQYLHGERIGVHICIRNNSNMVKKIYAMVOGVVDVLF	244
Db	192	LVSKGTFPSQGGKINLEVTLDREIYHGEKLAANVILNNRSKTYVKNIKVYVVOHCVEYVW	251
Qy	245	QNSSTRNVTYASLETSBEGCPIQPSSSLOKMYTLPLISSLNKKORCIGIALDQOIKRQDCLAS	304
Db	252	NAQSRHVALEFRREGCPITPGASPTKTFYVYLPPAASNNKORYIALDGYLKDODVNLAS	310
Qy	305	TYLLAQ PDQDADAFGVISAAYVULFLGALGGELSAELPFVLMHPPGPQ---	359
Db	311	STVSEGNKTTDAIGIVISYLRVKNLCNGTLLGGELQTDVDFPKLLHPAQTAREKQAQAI-	369
Qy	360	ADSOADVETPRD---TIDQOASVDEP 393	
Db	370	KKKKSIERTYRNSCYAADDDNDIVE 396	

RESULT 13
ARRB DROM

AC	STANDARD;	401 AA
P19107;	PRT;	
Q9VSN6;		
01-NOV-1999		

05-JUL-2004 (Rel. 16, Created)
25-JAN-2005 (Rel. 44, Last sequence update)
phosbrestin I (Arrestin B) (Arrestin update)
Name=Ar12; Synonym=ArB; (Arrestin 2) (49 kDa arrestin-like protein)
Drosophila melanogaster (Fruit fly)
Eukaryota; Metazoa; Arthropoda
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydraoidea; Drosophilidae; Drosophila
NCBI_Taxid=7227;

SEQUENCE FROM N.A.
MEDLINE=90232360; PubMed=2158671;
Yamada T., Takeuchi Y., Kmoroi N.,
Matsunoto H., Kobayashi H., Sakai Y.,
"A 49-kilodalton phosphoprotein in the *Drosophila*
arrestin homolog.",
Science 248:483-486(1990).
[2]

SEQUENCE FROM N.A.
STRAIN=Berkeley.

RA MEDLINE=20136006; PubMed=10731132; DOI=10.1126/science.287.5461.2185
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazet M.D., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Adair J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Bailew R.M., Baau A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkovec D., Bochan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.D., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibsygam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laemmel P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattedi B., McIntosh T.C., McLeod M.P., McHersom D.,
 RA Merkulov G., Milshina N.V., Nobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Palenart C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sinden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svaytekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT The genome sequence of *Drosophila melanogaster*.;
 RL Science 287:2185-2195 (2000).
 RN [3]
 RP GENOME REANNOTATION.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Malyburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,
 RA Betencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Ruso S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review".
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [4]
 RP PHOSPHORYLATION AND PROBABLE FUNCTION.
 RX MEDLINE=91282780; PubMed=1905538;
 RA Matsuno H., Yamada T.;
 RT "Phosretins I and II: arrestin homologs which undergo differential
 RT light-induced phosphorylation in the *Drosophila* photoreceptor in
 RT vivo".;
 RL Biochem. Biophys. Res. Commun. 177:1306-1312(1991).
 RN [5]
 RP PHOSPHORYLATION SITE SER-366.
 RX MEDLINE=9424441; PubMed=8185954; DOI=10.1016/0896-6273(94)90309-3;
 RA Matsuno H., Kurien B.T., Takagi Y., Kahn E.S., Kinumi T., Komori N.,
 RA Yamada T., Hayashi F., Isono K., Pak W.L.;
 RT "Phosretin I undergoes the earliest light-induced phosphorylation by
 RT a calcium/calmodulin-dependent protein kinase in *Drosophila*
 RT photoreceptors".;
 RL Neuron 12:997-1010(1994).
 CC -1- FUNCTION: Probably plays an important role in the photoreceptor
 CC transduction.
 CC -1- TISSUE SPECIFICITY: Inner and outer segments, and the inner
 CC plexiform regions of the retina.
 CC -1- PTM: Phosphorylated upon light exposure.
 CC -1- SIMILARITY: Belongs to the arrestin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.embnet.org/licenses/>
 CC or send an email to license@ebi-ebi.ch).
 CC -----
 DR EMBL; M32141; AAA28833.1; -;
 DR EMBL; A8003554; AAF50380.1; -;
 DR PIR; A34856; A34856.
 DR HSP; P17870; 1G4M.
 DR FlyBase; FBgn0000121; Arr2.
 DR GO; GO:0005624; C:membrane fraction; IDA.
 DR GO; GO:0016028; C:rhodome; IDA.
 DR GO; GO:0005625; C:soluble fraction; IDA.
 DR GO; GO:0016060; P:metaphodopsin inactivation; IMP.
 DR InterPro; IPR00698; Arrestin.
 DR InterPro; IPR011022; Arrestin_C.
 DR InterPro; IPR011021; Arrestin_N.
 DR Pfam; PF02752; Arrestin_C_1.
 DR Pfam; PF00339; Arrestin_N_1.
 DR PRINTS; PR00309; ARRESTIN.
 DR PRODOM; PD002099; Arrestin; 2.
 DR PROSITE; PS00295; ARRESTINS; 1.
 KW Phosphorylation; Sensory transduction; Vision.
 FT MOD RES 366
 FT VARIANT 109 109 N -> S.
 FT CONFLICT 111 111 Y -> H (in Ref. 1).
 SQ SEQUENCE 401 AA; 45028 MW; 12076E0DA8F0D87 CRC64;
 Query March 49.8%; Score 982.5; DB 1; Length 401;
 Best Local Similarity 48.7%; Pred. No. 8.3e-68;
 Matches 191; Conservative 73; Mismatches 115; Indels 13; Gaps 6;
 QY 1 MVYNFKVKKCAPNGKVTLYMGRDPYDVHSGVEPIDGIVLDEYIRDNKRFVGQIVCS 60
 DB 1 MVSVKVFKFATPVGKVTFLGRDRFDIDYCPVDGVIVEDYLUK-NRKVGQIATTT 59
 QY 61 FRYREDEWGLNFOKELCLASFOIYPRPEKSKQETQKQERLLKLGSAVIFTPNIS 120
 DB 60 YRYREDEWGLNFOKELCLASFOIYPRPEKSKQETQKQERLLKLGSAVIFTPNIS 118
 QY 121 PNPSSVTLQOGBDNDPCGVSYVYKIPAGESEDRTHRSVTYLGIRKIQAPRKQO 180
 DB 119 PNPSSVTLQOGBDNDPCGVSYVYKIPAGESEDRTHRSVTYLGIRKIQAPRKQO 178
 QY 181 Q-PTLYRKDFMLSPGELLEVTLKQLYLHGERIGVNICIRNNSNKKVKKIKAMVOGV 239
 DB 179 RLPSLSVSKGFTPSNGKISLEVTLDREIYHGEKTAFTVQVSNNSKKSIVKCFIVQHT 238
 QY 240 DVLVFGNGSRNRYVASLETBGCPIYOGSSIQKMYITPLLSNKKORGIALLDQIKROD 299
 DB 239 EITWV-NAQFSKVAQDETREGGCPITPGANITKTFYVILPLAANNKHGIALDHLDED 297
 QY 300 QCLASTTLAQ-PQGRAPFVITSYAVKVLFGALGSELASLPVLMHPKPGT----- 353
 DB 298 VNLASTVMOEGSGTGAAGVIVSYVRILKNCSTLGGEMOTDVFYKLIQAPAGTIEKKR 357
 QY 354 ---KAVYIHDSQADVFETFRDITDOQASVDF 382
 DB 358 SNMKKKMSIEQHRNVAGYVQDDDDNIVFEDF 389
 RESULT 14
 ARRB DROMI STANDARD; PRT; 401 AA.
 AC P19108;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Phosretin I (Arrestin B) (Arrestin 2) (49 kDa arrestin-like protein).
 GN Name=Arr2; Synonyms=ArrB;

OS Drosophila miranda (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydriidae; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7229;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S204;
 RX MEDLINE=91016944; PubMed=2216789;
 RA Krishnan R., Ganguly R.,
 RT "Nucleotide sequence of the arrestin-like 49 Kd protein gene of
 RL Drosophila miranda."
 CC Nucleic Acids Res. 18:5894-5894(1990).
 CC -1- FUNCTION: Undergoes light-induced phosphorylation, probably plays
 CC an important role in the photoreceptor transduction.
 CC -1- TISSUE SPECIFICITY: Inner and outer segments, and the inner
 CC -1- SIMILARITY: Belongs to the arrestin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X54064; CA38019.1; -
 DR PIR; S11566; S11566.
 DR HSSP; P17870; 1G4M.
 DR R1ybase; Pfam0012552; Dmlr\Arr2.
 DR InterPro; IPR000698; Arrestin.
 DR InterPro; IPR011022; Arrestin_C.
 DR Pfam; PF02752; Arrestin_N.
 DR Pfam; PF00339; Arrestin_N.1.
 DR PRINTS; PR00309; ARRESTIN.
 DR PRODOM; PD00209; ARRESTIN.
 DR PROSITE; PS00295; ARRESTIN; 1.
 DR Phosphorylation; Sensory transduction; Vision.
 KW KW
 QO SEQUENCE 401 AA; 45014 MW; E46EBB2485AF99C CRC64;
 Query Match 49.5%; Score 977.5; DB 1; Length 401;
 Best Local Similarity 48.5%; Pred. No. 2e-67;
 Matches 190; Conservative 76; Mismatches 113; Indels 13; Gaps 6;
 QY 1 M V N V F V F K K A P N G K V T L Y M G R D P V D V S G V E P I G V I V L D E Y I R D N R K V F G Q I V C S 60
 DB 1 M V S V A V F E K K A T P N G K V T F Y L G R D F I D H D Y C P V D G V I V E P E Y L K - N R K V F G Q L A T T 59
 QY 61 F Y Y G R E D E V M G I N F O K E L C L A S E Q I Y P R P E K S D K E Q T K L Q E R L K L G S N A I P F T P N I S 120
 DB 60 Y Y Y G R E D E V M G V F S K E L I L S R D E I V P M - T N P M E M T P T Q E K L V R K L G S N A H P F T H F P P 118
 QY 121 P A P S S V T L Q O G E D N D P C G V S Y V Y K I F A G E S E T D T H R S T V T L G I R K I Q A P T K Q Q 180
 DB 119 P N S P S S V T L Q O G E D N G D N K P L G V E Y T I R A F G S E D D R H K R S W S L V I K K I Q V A P L A N R Q 178
 QY 181 Q - P C T I V R K D F M L S P G E L E V T L D K O L Y L G E R I G V N I C I R N N S K M V K K I K A M V O G V D 239
 DB 179 R L P S S L V S K G F T P S N G K I S L E V T I L D R E I Y H G K V A T A V Q I N N S K S V S I K C F I Q H T 238
 QY 240 D V T L F O N G S Y N T V A S L E T S E G C P I Q P G S S L Q K W Y I L P T L S N K O R G I A L D G I K R O D 299
 DB 239 E I I M V - N A P S K A V A O L E T K E C P I T P G A N L K T F Y I L P L A A N N K D H G I A L D G I L K D B 297
 QY 300 Q C L A S T T L I A Q - P D O R D A F G V I S A V A V K L F G A L G E I S A E L P V L W H P K P G T - - - - - 353
 DB 298 V N L A S T I V M O E G K N G D A G I V I S V A I K N C G T L G E M Q T D V P K L Q P A P G T I E K R 357
 QY 354 - - - K A K V I H A D S Q A V E T F R O D T I D Q A S V D F 382
 DB 358 S V A M K M K S I E O H N N V K G Y I Q D D D D N I V F E D F 389

RESULT 15
 ARRB CALVI
 ID ARRB CALVI STANDARD; PRT; 400 AA.
 AC P51487;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Phosrestin I (Arrestin B) (Arrestin 2).
 GN Name=ARR2;
 OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 CC Calliphoridae; Calliphora.
 OK NCBI_TaxID=7373;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
 RC TISSUE=Retina;
 RX MEDLINE=95014564; PubMed=7929436;
 RA Planger A., Malicki D., Whitney M., Paulsen R.,
 RT "Mechanism of arrestin 2 function in rhodometric photoreceptors."
 RL J. Biol. Chem. 269:26969-26975(1994).
 CC -1- FUNCTION: Directly interacts with light-activated rhodopsin
 CC thereby activating the phosphorylation of metarhodopsin.
 CC the dephosphorylation of metarhodopsin.
 CC -1- SIMILARITY: Belongs to the arrestin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X79073; CA55673.1; -
 DR PIR; B55081; B55081.
 DR HSSP; P17870; 1G4M.
 DR InterPro; IPR000698; Arrestin.
 DR InterPro; IPR011022; Arrestin_C.
 DR InterPro; IPR011021; Arrestin_N.
 DR Pfam; PF02752; Arrestin_N.
 DR Pfam; PF00339; Arrestin_N.1.
 DR PRINTS; PR00309; ARRESTIN.
 DR PRODOM; PD00209; ARRESTIN.
 DR PROSITE; PS00295; ARRESTIN; 1.
 DR Direct protein sequencing; Phosphorylation; Sensory transduction;
 KW Vision.
 FT INIT MET
 QO SEQUENCE 400 AA; 44781 MW; DA786D7E9FFBA429 CRC64;
 Query Match 47.5%; Score 937.5; DB 1; Length 400;
 Best Local Similarity 47.6%; Pred. No. 2.6e-64;
 Matches 186; Conservative 70; Mismatches 122; Indels 13; Gaps 6;
 QY 2 V Y N K V F K K A P N G K V T L Y M G R D P V D V S G V E P I G V I V L D E Y I R D N R K V F G Q I V C S F 61
 DB 1 V S V A V F E K K A T P N G K V T F Y L G R D F I D H D Y I D P V D G V I V E P E Y L K - N R K V F A Q L A T T 59
 QY 62 R Y R E D E V M G I N F O K E L C L A S E Q I Y P R P E K S D K E Q T K L Q E R L K L G S N A I P F T P N I S 121
 DB 60 R Y G R E D E V M G V F S K E L I L S R D E I V P M - G N S N M E M T P T Q E K L V R K L G S N A H P F T H F P P 118
 QY 122 N A P S S V T L Q O G E D N D P C G V S Y V Y K I F A G E S E T D T H R S T V T L G I R K I Q A P T K Q Q 181
 DB 119 N P S S V T L Q O G E D D I G R P L G V E Y T I R A V A D S E D D R H K R S W S L V I K K I Q V A P L A N R Q 178
 QY 182 - P C T I V R K D F M L S P G E L E V T L D K O L Y L G E R I G V N I C I R N N S K M V K K I K A M V O G V D 240
 DB 179 L P S S L V S K G F T P S N G K I S L E V T I L D R E I Y H G K V A T A V Q I N N S K K A V N I K V I I Q H T E 238
 QY 241 V L F O N G S Y N T V A S L E T S E G C P I Q P G S S L Q K W Y I L P T L S N K O R G I A L D G I K R O D 300


```

Db      239  ITWV-NMQFSKHVAQLETKEGCPITPGANLSKTFYLIPLASNNKDRHIGIALDGHKXDEDV 297
Qy      301  CLASTTLLAQ-PDORDAFGVIIISYAVVKYKFLGALGELSHELFPVLMHPKPGT----- 353
Db      298  NLASSTWQDGKSTGDACGIVISYSVRIKLNCGLGGEIQTDPFKLLQAPAGSVEKRS 357
Qy      354  --KAKVIHADSQLADVETFRQDTIDQASVDF 382
Db      358  NAMKKMKSIHQHNTKGYQDDDDNIVFEDF 388

```

Search completed: February 9, 2005, 23:45:33
 Job time : 66 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:18:51 / Search time 165 Seconds
(without alignments)
897.753 Million cell updates/sec

Title: US-10-056-405-2

Perfect score: 1973
Sequence: 1 MYVNFKVFKKCAPNGKVTLY.....ADVEFRDRTIDQASVDFE 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1973	100.0	383	5	ABP52833 Anopheles
2	1973	100.0	383	4	ABR84476 Anopheles
3	1384	70.1	364	4	ABR61736 Drosophila
4	1384	70.1	364	8	ADQ89588 Antagonist
5	1023	51.8	398	4	ABR84484 Mosquito
6	1002	50.8	470	4	ABR58174 Drosophila
7	982.5	49.8	401	4	ABR61964 Anopheles
8	907.5	46.0	410	4	ABR67780 Anopheles
9	903.5	45.8	418	4	ABR67778 Anopheles
10	903.5	45.8	452	5	ABG69495 Rat bait
11	900.5	45.6	418	5	ABG70174 Human pro
12	900.5	45.6	418	8	ADQ89786 Antagonist
13	897.5	45.5	418	4	AAAG67779 Anopheles
14	890	45.1	369	4	ABG13574 Novel hum
15	885	44.9	382	4	AAAG67785 Anopheles
16	885	44.9	409	4	AAAG67781 Anopheles
17	885	44.9	409	7	ADRS8183 Human pro
18	885	44.9	409	7	ADRS8179 Human pro
19	885	44.9	409	7	ADN95251 Human BEC
20	880	44.6	409	4	AAAG67784 Anopheles
21	874	44.3	360	6	ADA00635 Human bet
22	874	44.3	495	6	ADA00637 Human bet
23	869.5	44.1	410	5	ABG69496 Rat bait
24	869.5	44.1	410	7	ADRS8177 Rat Prote
25	869.5	44.1	410	7	ADRS8181 Rat Prote

26	869.5	44.1	526	8	ADR23187 Yellow fl
27	869.5	44.1	526	8	ADR23188 Green fl
28	869.5	44.1	526	8	ADR23186 Yellow fl
29	865	43.8	454	8	ADN24049 Bacterial
30	780.5	39.6	388	4	AAAG67783 Anopheles
31	780	39.5	405	4	AAAG67783 Anopheles
32	779.5	39.5	388	4	AAAG67782 Anopheles
33	761	38.6	405	8	ADQ14327 Human ret
34	761	38.6	405	8	ADR41708 Protein s
35	661	33.5	468	4	AAAG79747 Human pro
36	516.5	26.2	182	7	ADD27419 Human adi
37	516.5	26.2	182	7	ADD27127 Human adi
38	516.5	26.2	187	6	ABU70623 Human adi
39	516.5	26.2	187	7	ADD27507 Human adi
40	497.5	25.2	180	6	ADA00634 Human bet
41	497.5	25.2	315	6	ADA00636 Human bet
42	482.5	24.5	182	7	ADD27197 Human adi
43	361.5	18.3	212	4	AAAG73899 Human col
44	351.5	17.8	456	4	ABR58949 Drosophila
45	173	8.8	65	4	ABR42223 Peptide #

ALIGNMENTS

RESULT 1
ABP52833
ID ABP52833 standard; protein: 383 AA.
XX
AC ABP52833;
XX
DT 01-NOV-2002 (first entry)
XX
DE Anopheles gambiae arresctin 1 protein SEQ ID NO:2.
XX
KW Anopheles gambiae; mosquito; olfactory gene; arresctin 1; pest control;
KM odourant receptor; olfaction.
XX
OS Anopheles gambiae.
XX
PN W0200259274-A2.
XX
PD 01-AUG-2002.
XX
PE 26-JAN-2002; 2002MO-US002549.
XX
PR 26-JAN-2001; 2001US-0264649P.
XX
PR 24-JAN-2002; 2002US-00056405.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
PI Zwiabel LJ;
XX
DR WPI: 2002-627421/67.
XX
DR N-PSDB; ABQ75102.
XX
PT New mosquito olfaction polypeptides and polynucleotides, useful for
PT mosquito management, i.e. controlling the pest and disease vectors, or
PT for identifying pest control agents.
XX
PS Claim 18; Fig 2; 96pp; English.
XX
CC The present invention describes a purified Anopheles gambiae olfaction
CC polypeptide comprising a 383, 394, 380, 411, 412, 391, 157 or 401 residue
CC amino acid sequence (see ABP52833 to ABP52840) (S1), a conservatively
CC modified amino acid sequence of them, or a sequence of (S1) with at least
CC 20 consecutive residues. Also described: (1) an isolated polynucleotide
CC comprising: (a) a nucleotide sequence encoding the purified Anopheles
CC gambiae olfaction polypeptide; or (b) a nucleotide sequence that
CC hybridizes under stringent conditions to a hybridisation probe comprising
CC a 1964, 1239, 1142, 1236, 1194, 1176, 474 or 1206 nucleotide sequence
CC (see ABQ75102 to ABQ75105 and ABQ75110 to ABQ75113) (S2), or its
CC complement; and (2) a method for identifying an agent that binds to

CC mosquito olfaction molecules comprising: (a) providing an isolated:
CC mosquito olfaction molecule; (b) contacting a test agent with the
CC isolated mosquito olfaction molecule; and (c) detecting specific binding
CC of the test agent to the isolated mosquito olfaction molecule, where the
CC olfaction-binding identifies the test agent as a mosquito
CC method for mosquito management. The mosquito olfaction molecules are useful
CC modulate arrestin-odorant receptor interaction for substances that
CC arrestin 1 from the present invention

Sequence 383 AA;

Query Match
Best Local Similarity 100.0%; Score 1973; DB 5; Length 383;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYNFKVFKKCAKNGAVTLYMGRDPVHVGVEPIIDGIVLDDYIRNKRKVFQIVCS 60
DB 1 MYNFKVFKKCAKNGAVTLYMGRDPVHVGVEPIIDGIVLDDYIRNKRKVFQIVCS 60
QY 61 FRYGREDEVMGLNFOKEICLASBOIYPREKSDKEQOTLQERLLKLGSAIPFTFNIS 120
DB 61 FRYGREDEVMGLNFOKEICLASBOIYPREKSDKEQOTLQERLLKLGSAIPFTFNIS 120
QY 121 PNAISSVTLOQGEDNDGDPGCVSYVYKIFAGESETDTRHRSVTYTGIRKIOPAPTKQG 180
DB 121 PNAISSVTLOQGEDNDGDPGCVSYVYKIFAGESETDTRHRSVTYTGIRKIOPAPTKQG 180
QY 181 QPCTLVKRDPMLSFGLELEVTLDKQLYLHGERIGVNICIRNNNSNMVYKIKAMVQGV 240
DB 181 QPCTLVKRDPMLSFGLELEVTLDKQLYLHGERIGVNICIRNNNSNMVYKIKAMVQGV 240
QY 241 VVLFONGSYRNTVASLETSSECPIDPSSLOKMYITPLSSNKGRIALDQIIRQDQ 300
DB 241 VVLFONGSYRNTVASLETSSECPIDPSSLOKMYITPLSSNKGRIALDQIIRQDQ 300
QY 301 CLASTTLLAOPDQDAFGVLIISYAVVKLFALGELSAELPFVLMHPRPKTKAKVHA 360
DB 301 CLASTTLLAOPDQDAFGVLIISYAVVKLFALGELSAELPFVLMHPRPKTKAKVHA 360
QY 361 DSQADVETFRQDTIIDQASVDPE 383
DB 361 DSQADVETFRQDTIIDQASVDPE 383

RESULT 2
ID ABR84476 standard; protein; 383 AA.
AC ABR84476;
DT 15-JAN-2004 (first entry)
XX Mosquito olfaction molecule, arrestin 1.
XX Mosquito olfaction molecule, arrestin 1.
XX Anopheles gambiae.
XX WO2003076590-A2.
XX 18-SEP-2003.
XX 10-MAR-2003; 2003WO-US007174.
XX 08-MAR-2002; 2002US-00094240.
XX (UYVA-) UNIV VANDERBILT.
XX Zwiebel LJ;
XX

DR WPI; 2003-722331/68.
DR N-PSDB; ACF79716.
XX
PT New mosquito arrestin 1 and 2 genes and polypeptides, useful for
PT identifying mosquito olfaction molecule binding compounds which reduce
XX the ability of mosquitoes to locate sources of bloodmeal, e.g. humans.
PS Claim 1, Fig 2; 101pp; English.
XX
CC The present sequence is the protein sequence of arrestin 1, a novel
CC mosquito olfaction molecule. The invention provides 9 novel mosquito
CC polypeptides and the nucleic acids encoding them. These are arrestin
CC in a ligand-induced signal transduction pathway for the activation of
CC mosquito olfaction. Arrestin functions to inhibit the activated signal
CC transduction cascade. Thus, the odorant receptors act as an 'on' switch
CC and arrestin as an 'off' switch for the odorant detection system of the
CC mosquito. Methods are provided for identifying compounds that interfere
CC with the operation of the mosquito olfactory system, particularly the
CC control of mosquitoes, particularly by reducing their ability to locate
CC sources of bloodmeal

Sequence 383 AA;
Query Match
Best Local Similarity 100.0%; Score 1973; DB 7; Length 383;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYNFKVFKKCAKNGAVTLYMGRDPVHVGVEPIIDGIVLDDYIRNKRKVFQIVCS 60
DB 1 MYNFKVFKKCAKNGAVTLYMGRDPVHVGVEPIIDGIVLDDYIRNKRKVFQIVCS 60
QY 61 FRYGREDEVMGLNFOKEICLASBOIYPREKSDKEQOTLQERLLKLGSAIPFTFNIS 120
DB 61 FRYGREDEVMGLNFOKEICLASBOIYPREKSDKEQOTLQERLLKLGSAIPFTFNIS 120
QY 121 PNAISSVTLOQGEDNDGDPGCVSYVYKIFAGESETDTRHRSVTYTGIRKIOPAPTKQG 180
DB 121 PNAISSVTLOQGEDNDGDPGCVSYVYKIFAGESETDTRHRSVTYTGIRKIOPAPTKQG 180
QY 181 QPCTLVKRDPMLSFGLELEVTLDKQLYLHGERIGVNICIRNNNSNMVYKIKAMVQGV 240
DB 181 QPCTLVKRDPMLSFGLELEVTLDKQLYLHGERIGVNICIRNNNSNMVYKIKAMVQGV 240
QY 241 VVLFONGSYRNTVASLETSSECPIDPSSLOKMYITPLSSNKGRIALDQIIRQDQ 300
DB 241 VVLFONGSYRNTVASLETSSECPIDPSSLOKMYITPLSSNKGRIALDQIIRQDQ 300
QY 301 CLASTTLLAOPDQDAFGVLIISYAVVKLFALGELSAELPFVLMHPRPKTKAKVHA 360
DB 301 CLASTTLLAOPDQDAFGVLIISYAVVKLFALGELSAELPFVLMHPRPKTKAKVHA 360
QY 361 DSQADVETFRQDTIIDQASVDPE 383
DB 361 DSQADVETFRQDTIIDQASVDPE 383

RESULT 3
ID ABB61736 standard; protein; 364 AA.
AC ABB61736;
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 12000.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX

PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL05839.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 12000; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIFO at ftp.wifo.int/pub/published_pct_sequences
 XX
 SQ Sequence 364 AA;
 Query Match 70.1%; Score 1384; DB 4; Length 364;
 Best Local Similarity 70.3%; Pred. No. 1.5e-135;
 Matches 258; Conservative 51; Mismatches 54; Indels 4; Gaps 2;

QY 1 MVNFKYFKKCAENGKATLYMGKDFVDHVSQVEPIGIVLDDDEYRDNKRVGQIVCS 60
 DB 1 MVNFKYFKKCSFNNMTLYMNRDPVDSVTQVEPIGIIVLDDDEYRDNKRVGQIVCS 60
 QY 61 FRYGREDEVMGKDFVDSQVCP-PQKODIQLKQERLLKLGSAVFPVWMP 119
 DB 61 FRYGREDEVMGKDFVDSQVCP-PQKODIQLKQERLLKLGSAVFPVWMP 119
 QY 121 PNAFSSVTLQOGEEDNDPCGVSYVYKIFAGESETRTHRSVTTLGIRKIQFAPTRQOQ 180
 DB 120 PSSPASVVLQOKASDESGQVGVYFKIFGDSDCDRSHRSTINLGIRKQVAPTRQGI 179
 QY 181 QPCTLVKRDPMLSGELLEVTLDDKQLYHGERIGVNICIRNNSNKKVKKIKAVVOGVD 240
 DB 180 QPCTVVRKDFLISGELLEVTLDDKQLYHGERIGVNICIRNNSNKKVKKIKAVVOGVD 239
 QY 241 VVLFQNSYRNTVAISLTSBSCPIOPGSSLOKMWYLPPLSSNQKRGIALDGOIKRPOQ 300
 DB 240 VVLFQNSYRNTVAISLTSBSCPIOPGSSLOKMWYLPPLSSNQKRGIALDGOIKRPOQ 299
 QY 301 CLASTTLAOPDQDAFGVLIISYAVKYLFLGALGELSABLPLVLMHPKPGTAKVYIHA 360
 DB 300 ALASTTLIASGDADAFGIIVSYAVKYLFLGALGELSABLPLVLMHPKPGTAKVYIHA 356
 QY 361 DSGADVE 367
 DB 357 EAEGSIE 363
 RESULT 4
 ID ADO89588 standard; protein; 364 AA.
 XX
 AC ADO89588;

XX
 DT 21-OCT-2004 (first entry)
 XX
 DB Antagonist of cell cycle progression polypeptide #9.
 XX
 KW Cyostatic; cancer; cell division cycle; mitosis; meiosis;
 KW cell cycle progression.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO2004063362-A2.
 XX
 PD 29-JUL-2004.
 XX
 PR 31-DEC-2003; 2003WO-GB005635.
 XX
 PR 10-JAN-2003; 2003US-0439123P.
 PR 06-MAY-2003; 2003US-0468402P.
 XX
 PA (CYCL-) CYCLACEL LTD.
 XX
 PI Glover D, Bell G, Frenz L, Midgley C;
 XX
 DR WPI; 2004-544089/52.
 DR N-PSDB; ADO89587.
 XX
 PT New cell cycle progression genes and proteins for modulating cell cycle
 PT progression in cells, for preventing, creating or diagnosing cell
 PT proliferative diseases (e.g. cancer) or for identifying modulators of
 PT mitosis or meiosis.
 XX
 PS Claim 2; SEQ ID NO 18; 461pp; English.
 XX
 CC The present invention relates to a polynucleotide for preventing,
 CC treating or diagnosing a disease in an individual. The composition or the
 CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for
 CC diagnosing, preventing or treating diseases (e.g. cell proliferative
 CC diseases such as cancer) in an individual. These may also be used for
 CC identifying substances capable of binding to or modulating the function
 CC of the polypeptide, capable of affecting the function of the
 CC corresponding gene, or capable of inhibiting the cell division cycle or
 CC cell cycle progression, preferably mitosis and/or meiosis. The present
 CC sequence represents an antagonist of cell cycle progression protein
 CC sequence.

SQ Sequence 364 AA;
 Query Match 70.1%; Score 1384; DB 8; Length 364;
 Best Local Similarity 70.3%; Pred. No. 1.5e-135;
 Matches 258; Conservative 51; Mismatches 54; Indels 4; Gaps 2;
 QY 1 MVNFKYFKKCAENGKATLYMGKDFVDHVSQVEPIGIVLDDDEYRDNKRVGQIVCS 60
 DB 1 MVNFKYFKKCSFNNMTLYMNRDPVDSVTQVEPIGIIVLDDDEYRDNKRVGQIVCS 60
 QY 61 FRYGREDEVMGKDFVDSQVCP-PQKODIQLKQERLLKLGSAVFPVWMP 119
 DB 61 FRYGREDEVMGKDFVDSQVCP-PQKODIQLKQERLLKLGSAVFPVWMP 119
 QY 121 PNAFSSVTLQOGEEDNDPCGVSYVYKIFAGESETRTHRSVTTLGIRKIQFAPTRQOQ 180
 DB 120 PSSPASVVLQOKASDESGQVGVYFKIFGDSDCDRSHRSTINLGIRKQVAPTRQGI 179
 QY 181 QPCTLVKRDPMLSGELLEVTLDDKQLYHGERIGVNICIRNNSNKKVKKIKAVVOGVD 240
 DB 180 QPCTVVRKDFLISGELLEVTLDDKQLYHGERIGVNICIRNNSNKKVKKIKAVVOGVD 239
 QY 241 VVLFQNSYRNTVAISLTSBSCPIOPGSSLOKMWYLPPLSSNQKRGIALDGOIKRPOQ 300
 DB 240 VVLFQNSYRNTVAISLTSBSCPIOPGSSLOKMWYLPPLSSNQKRGIALDGOIKRPOQ 299
 QY 301 CLASTTLAOPDQDAFGVLIISYAVKYLFLGALGELSABLPLVLMHPKPGTAKVYIHA 360
 DB 300 ALASTTLIASGDADAFGIIVSYAVKYLFLGALGELSABLPLVLMHPKPGTAKVYIHA 356

Db 300 ALASTILIASQDARDAFGIIVAVKVLFLGALGSELCAELPFLIMPKPSRKAQL--- 356
 QY 361 DSQADVE 367
 : : : :
 Db 357 EAEGSIE 363

RESULT 5
 ABR84484
 ID ABR84484 standard; protein; 398 AA.
 AC ABR84484;
 XX
 DT 15-JAN-2004 (first entry)
 DE Mosquito olfaction molecule, arrestin 2.
 XX
 KW Arrestin 2; mosquito; olfaction; insecticide; antimalarial.
 OS Anopheles gambiae.
 XX
 FH Key
 FT Location/Qualifiers
 FT Misc-difference 152
 XX /note= "Encoded by AG"
 PN MO2003076590-A2.
 XX
 PD 18-SEP-2003.
 XX
 PF 10-MAR-2003; 2003WO-US007174.
 XX
 PR 08-MAR-2002; 2002US-00094240.
 XX
 PA (UYVA-) UNIV VANDERBILT.
 XX
 PI Zwiebel LJ;
 DR WPI; 2003-722331/68.
 DR N-PSDB; ACF79731.
 XX
 PT New mosquito arrestin 1 and 2 genes and polypeptides, useful for
 XX identifying mosquito olfaction molecule binding compounds which reduce
 PS the ability of mosquitoes to locate sources of bloodmeal, e.g. humans.
 XX
 XX Claim 16; Fig 12b; 101pp; English.
 CC The present sequence is the protein sequence of arrestin 2, a novel
 CC mosquito olfaction molecule. The invention provides 9 novel mosquito
 CC and 2 odorant receptor molecules encoding them. These are arrestins 1
 CC in a ligand-induced signal transduction pathway for the activation of
 CC mosquito olfaction. Arrestin functions to inhibit the activation of
 CC transduction cascade. Thus, the odorant receptors act as an 'on' switch,
 CC and arrestin as an 'off' switch for the odorant detection signal.
 CC Mosquito. Methods are provided for identifying compounds that interfere
 CC with the operation of the mosquito olfactory system, particularly the
 CC compounds that modulate arrestin 2 activity. These are useful for the
 CC control of mosquitoes, particularly by reducing their ability to locate
 CC sources of bloodmeal.
 XX
 XX Sequence 398 AA:

Query Match
 Best Local Similarity 51.8%; Score 1023; DB 7; Length 398;
 Matches 204; Conservative 65; Mismatches 111; Indels 10; Gaps 7;
 QY 1 MYNFEVFKKCAPNKVTLVNGKRDVGVPEPIDGVILDDYITDNKRVGQIVCS 60
 Db 1 MYVAAYVFKKAPNGKLTIVLGRKDPIDHDYCDPIDGVILDEBYLR-GRKVFQGLITT 59
 QY 61 FRYGEEDEWVGLMFOELCLASEQIYPRPEKSDGEQKLOERILKLKGSNAIPFENIS 120
 Db 60 YRGRSEDEWVGKVFYSKENVLTKEQIYPM-ENAMNEMTQWERLVKLGANAPFTTHRP 118

QY 121 PNAPSSVTLQOGEEDNDPCGVSYVYKIFAGESETDRTHRRSTVTLGIRKIOPAPTKQO 180
 Db 119 SWAPSSVTLQAGEDDDGKPLGVEYAIKNAVGEDESDGKHRSAAVTLTIKLGQVAPVSKR 178
 QY 181 Q-PCITVRKQPMLEPELEETLDRKQYLHGERIGVNICIRNNSKMKIKIKANVQGV 239
 Db 179 RLPSLSVSKGFTFSSQCKINLETTLDRETYHGEKIANIYVTNNSKTKYKSTKCFVQHC 238
 QY 240 DVVLPONGSYRNTVASLETSEGCPIQGSLSLQVMTLTPLSNNQORGIALDQIIRPO 299
 Db 239 EYTMV-NAQFSKHLASLETREGCITFGASFTKSPFLVPLASSNDRRGIALDGHLEKD 297
 QY 300 OCLASTTLIAQPD-QRAFGVITISYAVKVLFLGALGSELCAELPFLIMPKRPGT--KAK 356
 Db 298 VNLASTLISGKCPSPDMGIVISYELRVKLNCGTLGSELQTDVDFPLNPPAPSVVERK 357
 QY 357 VTHADSQADVEYTRQDPT--IDQASVDEP 383
 Db 358 VNAKKKKKSTERRHRENSHYADDDNIVFE 387

RESULT 6
 ABB58174
 ID ABB58174 standard; protein; 470 AA.
 AC ABB58174;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 1314.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN MO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PA 11-JUL-2000; 2000US-00614150.
 XX
 PI (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW,
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02277.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signaling and cell-cell
 XX interactions.
 PS Disclosure; SEQ ID NO 1314; 21pp + Sequence listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development and
 CC insecticides, therapeutics and pharmaceutical drugs. The invention of
 CC sequences genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 470 AA:

Query Match
 50.8%; Score 1002; DB 4; Length 470.

DR WPI: 2001-602637/68.

XX Identifying a G protein-coupled receptor ligand, useful for treating e.g.
PT retinitis pigmentosa, color blindness or neurological disorders, uses
PT phosphorylation-independent arrestin mutants particularly suited for in
XX vitro screening assays.

XX Disclosure: Page 46; 47pp; English.

CC The present sequence represents an arrestin protein. The specification
CC describes phosphorylation-independent arrestin mutants. These mutants are
CC used in screening assays to identify ligands and/or modulators of G
CC protein-coupled receptors (GPCRs). A method for identifying a GPCR ligand
CC comprises contacting a composition comprising the GPCR and a
CC constitutively active arrestin mutant with a test compound; and
CC determining the ability of the test compound to modulate binding of
CC the test compound to the GPCR, where modulation of binding indicates that
CC potential ligands and/or modulators of GPCRs, particularly modulators,
CC for use as human therapeutics. Modulators of GPCR may be used for
CC treating patients having e.g. retinitis pigmentosa, stationary night
CC blindness, colour blindness, nephrogenic DI, isolated glucocorticoid
CC deficiency, hyperparathyroidism and neurological disorders. The
CC small molecule library of screening pluralities of test compounds (e.g. a
CC plurality of GPCRs. The methods may be used in screening assays for
CC identification of natural and surrogate agonists of orphan GPCRs, and for
CC identification of GPCR antagonists and/or agonists

Sequence 410 AA:

Query Match 46.0%; Score 907.5; DB 4; Length 410;
Best Local Similarity 47.2%; Pred. No. 1.5e-85;
Matches 176; Conservative 78; Mismatches 112; Indels 7; Gaps 4;

```

QY 6 KVFKKCAPNGKVTLYMGKRDVDSVGVPIIDGIVLDDERYIDNRKVGQIVCSFRYGR 65
DB 7 RVFKKASPVGKLTLYVIGKDFVDHIDVDPDGVAVDEYLAKE-RRVYVTLTCAFRGR 65
QY 66 EEDVWGLNFQKELCLASEQIYPRPKSDKEQTKLQERLLKLGNSALPFTFNISPAVS 125
DB 66 EEDVWGLNFQKELCLASEQIYPRPKSDKEQTKLQERLLKLGNSALPFTFNISPAVS 125
QY 126 SVTLQGEDNDGDPGCVSVYVYKIFAGESETDRTSRSTVTLGIRKIQAPTKQGOQPC 125
DB 126 SVTLQGEDNDGDPGCVSVYVYKIFAGESETDRTSRSTVTLGIRKIQAPTKQGOQPC 125
QY 186 VRDPMLSPEELEVTLLDKQLYLHGERIGNICIRNNSNMVKIKAMVQGVDPVLFQ 185
DB 186 VRDPMLSPEELEVTLLDKQLYLHGERIGNICIRNNSNMVKIKAMVQGVDPVLFQ 185
QY 186 TTRQFLMSDKPLHLEASLDKEIYNGEPISVAVHNTNNTKVKIKISVQYALICLPN 245
DB 186 TTRQFLMSDKPLHLEASLDKEIYNGEPISVAVHNTNNTKVKIKISVQYALICLPN 245
QY 246 NSGYRNTVASLETSGCPRIQSGSLQKVMYITPLLSNKRQRIALDQIKQDQCLAST 305
DB 246 NSGYRNTVASLETSGCPRIQSGSLQKVMYITPLLSNKRQRIALDQIKQDQCLAST 305
QY 246 TAQYKCPVMEAD--IVAPSSFTFCVYVTLTFLANNREKGLALDGLKHEIDTNLASS 303
DB 246 TAQYKCPVMEAD--IVAPSSFTFCVYVTLTFLANNREKGLALDGLKHEIDTNLASS 303
QY 306 TLLAQRDQDAFGYIISYAVKVLFLGALGSELSALPLPVLMHPKGRK--AKVIHADS 362
DB 306 TLLAQRDQDAFGYIISYAVKVLFLGALGSELSALPLPVLMHPKGRK--AKVIHADS 362
QY 363 QADVETFRQDTID 375
DB 363 QADVETFRQDTID 375
QY 363 PVDTNLIETDND 375
DB 363 PVDTNLIETDND 375

```

RESULT 9
AAG67778
ID AAG67778 standard; protein; 418 AA.
XX AAG67778;
XX AAG67778;
XX AAG67778;
DT 10-DEC-2001 (first entry)
XX

DE Amino acid sequence of bovine beta-arrestin 1A.
XX
XX Arrestin; phosphorylation-independent arrestin mutant;
KW G protein-coupled receptor; GPCR; GPCR ligand; retinitis pigmentosa;
KW stationary night blindness; colour blindness; nephrogenic DI;
KW isolated glucocorticoid deficiency; hyperfunctioning thyroid adenoma;
KW familial hypocalcaemic hypercalcaemia; hyperparathyroidism;
XX neurological disorder.

Bos sp.
XX
XX Key Location/Qualifiers
XX Region 155..184
XX /note="putative phosphorylation-recognition region"

MO200167106-A2.
XX
XX 13-SEP-2001.

05-MAR-2001; 2001MO-US007304.
XX
XX 03-MAR-2000; 2000US-0186706P.

(MILL-) MILLENNIUM PHARM INC.
XX
XX Berstein G;
XX
XX WPI: 2001-602637/68.

PT Identifying a G protein-coupled receptor ligand, useful for treating e.g.
PT retinitis pigmentosa, color blindness or neurological disorders, uses
PT phosphorylation-independent arrestin mutants particularly suited for in
XX vitro screening assays.

XX Disclosure: Page 46; 47pp; English.

CC The present sequence represents an arrestin protein. The specification
CC describes phosphorylation-independent arrestin mutants. These mutants are
CC used in screening assays to identify ligands and/or modulators of G
CC protein-coupled receptors (GPCRs). A method for identifying a GPCR ligand
CC comprises contacting a composition comprising the GPCR and a
CC constitutively active arrestin mutant with a test compound; and
CC determining the ability of the test compound to modulate binding of
CC the test compound to the GPCR, where modulation of binding indicates that
CC potential ligands and/or modulators of GPCRs, particularly modulators,
CC for use as human therapeutics. Modulators of GPCR may be used for
CC treating patients having e.g. retinitis pigmentosa, stationary night
CC blindness, colour blindness, nephrogenic DI, isolated glucocorticoid
CC deficiency, hyperparathyroidism and neurological disorders. The
CC small molecule library of screening pluralities of test compounds (e.g. a
CC plurality of GPCRs. The methods may be used in screening assays for
CC identification of natural and surrogate agonists of orphan GPCRs, and for
CC identification of GPCR antagonists and/or agonists

Sequence 418 AA:

Query Match 45.8%; Score 903.5; DB 4; Length 418;
Best Local Similarity 47.1%; Pred. No. 4.2e-85;
Matches 179; Conservative 76; Mismatches 112; Indels 13; Gaps 5;

```

QY 6 KVFKKCAPNGKVTLYMGKRDVDSVGVPIIDGIVLDDERYIDNRKVGQIVCSFRYGR 65
DB 7 RVFKKASPVGKLTLYVIGKDFVDHIDVDPDGVAVDEYLAKE-RRVYVTLTCAFRGR 65
QY 66 EEDVWGLNFQKELCLASEQIYPRPKSDKEQTKLQERLLKLGNSALPFTFNISPAVS 125
DB 66 EEDVWGLNFQKELCLASEQIYPRPKSDKEQTKLQERLLKLGNSALPFTFNISPAVS 125
QY 126 SVTLQGEDNDGDPGCVSVYVYKIFAGESETDRTSRSTVTLGIRKIQAPTKQGOQPC 125
DB 126 SVTLQGEDNDGDPGCVSVYVYKIFAGESETDRTSRSTVTLGIRKIQAPTKQGOQPC 125

```


Db 126 SVTLQPEPBDTGKACGVDEYKAFCAENLEEKIKHRSVRLVRKQYAPERPPOPTAE 185
 Qy 186 VRKDFMLSPGELTEVTLDKQLYLHGERIGVNICIRNNSNMVKKIKAMQOGVDVLFQ 245
 Db 186 TTRQFLMSDKPLHLEASLDKEIYHGERISVNVHTNNTKTVKKIKISVQYADICLFN 245
 Qy 246 NGSYRNTVASLETSEGCPIOPGSSLOKVMYLTPLLSNKKORGIALDQIKRODOCLAST 305
 Db 246 TAQYKCPVAMEADD--TVAPSSIFCKVYTLTPFLANNREKRGALDGLKHEDTNLASS 303
 Qy 306 TLLAQPDQDAFGVIISYAVKVLFL--GALGELSA-----ELPVLMPKPGCTK--A 355
 Db 304 TLLREGANREILGIIVSKVKVLVSRGGLDGLASSDVAVELPFTLMHPKPEBPPHR 363
 Qy 356 KVIHADSQADVETFRDPTID 375
 Db 364 EVPEHETPVDNLIELDTND 383
 RESULT 10
 ABG69495
 ID ABG69495 standard; protein, 452 AA.
 XX
 AC ABG69495;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Rat bait protein beta arrestin 1.
 XX
 KM Rat; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
 KM non-insulin diabetes mellitus; obesity; selected interacting domain; SID;
 KM protein-protein interaction map; PIM; anorectic; metabolic disorder.
 XX
 OS Rattus sp.
 XX
 PN WO200253726-A2.
 XX
 PD 11-JUL-2002.
 XX
 PF 28-DEC-2001; 2001WO-EP015423.
 XX
 PR 02-JAN-2001; 2001US-0259377P.
 XX
 PA (HYBR-) HYBRIGENICS.
 XX (CNRS) CENT NAT RECH SCT.
 XX
 PI Legrain P, Marullo S, Jockers R;
 XX
 DR WPI; 2002-583612/62.
 XX N-PSDB; ABS51031.
 PT Novel complex of protein-protein interactions in adipocyte cells for
 PT identifying compounds that modulate the protein-protein interactions and
 PT useful for treating obesity and metabolic disorders.
 XX
 PS Claim 1; Page 52; 125pp; English.
 CC The invention relates to a complex of protein-protein interactions
 CC (forming a protein-protein interaction map, PIM) in adipocyte cells as
 CC defined in the specification, or polynucleotides in adipocytes encoding
 CC for the polypeptides. Also included are a recombinant cell expressing the
 CC interacting polypeptides and a method of selecting a modulating compound
 CC in adipocyte cells, by cultivating a recombinant host cell on a selective
 CC medium containing a modulating compound and a reporter gene the
 CC expression of which is toxic for the recombinant host cell which is
 CC transformed with two vectors, where the first vector comprises a
 CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain
 CC and the second vector comprising a polynucleotide encoding a second
 CC hybrid polypeptide and an activating domain that activates the toxic
 CC reporter gene, when the first and second hybrid polypeptides interact and
 CC selecting the modulating compound which inhibits the growth of the
 CC recombinant host cell (i.e. using the yeast two-hybrid system). The
 CC complexes are useful for identifying compounds that modulate the protein-

CC protein interactions and useful for treating obesity and metabolic
 CC disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The
 CC compound isolated by the method is useful for treating and preventing
 CC obesity or metabolic diseases. The interactions between the proteins of
 CC the complex further define a set of selected interacting domains, SID.
 CC The present sequence represents a member of the protein complex of the
 CC invention, used as the bait protein in the yeast two- hybrid assay
 CC
 SO Sequence 452 AA;
 Query Match 45.8%; Score 903.5; DB 5; Length 452;
 Best Local Similarity 47.1%; Pred. No. 4,7e-85;
 Matches 179; Conservative 76; Mismatches 112; Indels 13; Gaps 5;
 Qy 6 KVFKKCAPNGKRVTLVMGKRDFVDHVSGEPLIDGIVLDDRYIRNKRKVFQGVSPRYGR 65
 Db 7 RVFKKASPNKGLTVYLGKDFVDHIDLVDPVDGVLVDPETLKE--RVYVYTLTCAFRGR 65
 Qy 66 EEDVEMGLNFQKEICLASQIYPRPEKSKDEQTYLQERLYKKJGSNA1PFTFNISPNAPS 125
 Db 66 EDDLVLGLTRKOLFVANNQSPFPAPEDKKPLRLQERLYKKJGHNHYPTFELIPNLPIC 125
 Qy 126 SVTLQGEDNDGDPGVSYVYKIPAGESETRDRTHRSVTVLGIRKIQAPATKOGQOPCTL 185
 Db 126 SVTLQPEPBDTGKACGVDEYKAFCAENLEEKIKHRSVRLVRKQYAPERPPOPTAE 185
 Qy 186 VRKDFMLSPGELTEVTLDKQLYLHGERIGVNICIRNNSNMVKKIKAMQOGVDVLFQ 245
 Db 186 TTRQFLMSDKPLHLEASLDKEIYHGERISVNVHTNNTKTVKKIKISVQYADICLFN 245
 Qy 246 NGSYRNTVASLETSEGCPIOPGSSLOKVMYLTPLLSNKKORGIALDQIKRODOCLAST 305
 Db 246 TAQYKCPVAMEADD--TVAPSSIFCKVYTLTPFLANNREKRGALDGLKHEDTNLASS 303
 Qy 306 TLLAQPDQDAFGVIISYAVKVLFL--GALGELSA-----ELPVLMPKPGCTK--A 355
 Db 304 TLLREGANREILGIIVSKVKVLVSRGGLDGLASSDVAVELPFTLMHPKPEBPPHR 363
 Qy 356 KVIHADSQADVETFRDPTID 375
 Db 364 EVPEHETPVDNLIELDTND 383
 RESULT 11
 ABG70174
 ID ABG70174 standard; protein, 418 AA.
 XX
 AC ABG70174;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human prey protein for Shigella flexner 9 #4.
 XX
 KM Prey protein; ospb; ospd1; ipad; ipac; ipan9; 8; ospg; ospc1; Shigella;
 KM shigellosis; bacillary dysentery; antibacterial; yeast two-hybrid system;
 KM protein-protein interaction; SID; selected interacting domain; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200257303-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EP000777.
 XX
 PR 12-JAN-2001; 2001US-0261130P.
 XX
 PA (HYBR-) HYBRIGENICS.
 XX
 PI Legrain P;
 XX
 DR WPI; 2002-599706/64.
 DR N-PSDB; ABS51567.

XX	RESULT 13
XX	AAG67779
ID	AAG67779 standard; protein; 418 AA.
AC	
XX	AAG67779;
DT	10-DEC-2001 (first entry)
DE	Amino acid sequence of human beta-arrestin 1A.
XX	
KW	Arrestin; phosphorylation-independent arrestin mutant;
KM	G protein-coupled receptor; GPCR; GPCR ligand; retinitis pigmentosa;
KM	stationary night blindness; colour blindness; nephrogenic DI;
KM	isolated glucocorticoid deficiency; hyperfunctioning thyroid adenoma;
KM	familial hypocalcemic hypercalcemia; hyperparathyroidism;
XX	neurological disorder.
OS	
XX	Homo sapiens.
FH	
FT	Key Location/Qualifiers
FT	Region 155..184
FT	/note= "putative phosphorylation-recognition region"
XX	
PN	MOZ00167106-A2.
XX	
PD	13-SEP-2001.
PF	05-MAR-2001; 2001WO-US007304.
XX	
XX	03-MAR-2000; 2000US-0186706P.
PA	(MLL-) MULLENMUM PHARM INC.
P1	Berstein G;
XX	
DR	WPI; 2001-602637/68.
XX	
PT	Identifying a G protein-coupled receptor ligand, useful for treating e.g.
PT	retinitis pigmentosa, color blindness or neurological disorders, uses
PT	phosphorylation-independent arrestin mutants particularly suited for in
PT	vitro screening assays.
XX	
PS	Disclosure; Page 46; 47pp; English.
XX	
CC	The present sequence represents an arrestin protein. The specification
CC	describes phosphorylation-independent arrestin mutants. These mutants are
CC	used in screening assays to identify ligands and/or modulators of G
CC	protein-coupled receptors (GPCRs). A method for identifying a GPCR ligand
CC	comprises contacting a composition comprising the GPCR and a
CC	constitutively active arrestin mutant with a test compound; and
CC	determining the ability of the test compound to modulate binding of
CC	arrestin mutant to the GPCR, where modulation of binding indicates that
CC	the test compound is a GPCR ligand. The method is used for identifying
CC	potential ligands and/or modulators of GPCRs, particularly modulators,
CC	for use as human therapeutics. Modulators of GPCR may be used for
CC	treating patients having e.g. retinitis pigmentosa, stationary night
CC	blindness, colour blindness, nephrogenic DI, isolated glucocorticoid
CC	deficiency, hyperfunctioning thyroid adenomas, familial hypocalcaemic
CC	hypercalcaemia, hyperparathyroidism and neurological disorders. The
CC	method may be used for screening pluralities of test compounds (e.g. a
CC	small molecule library of compounds) or a composition containing a
CC	plurality of GPCRs. The methods may be used in screening assays for
CC	identification of natural and surrogate agonists of orphan GPCRs, and for
CC	identification of GPCR antagonists and/or agonists
SQ	
Sequence	418 AA;
Query Match	45.5%; Score 897.5; DB 4; Length 418;
Best Local Similarity	46.8%; Pred.No.1,8e-84;
Matches 178; Conservative	76; Mismatches 113; Indels 13; Gaps 5

[illegible]

CC	ABG13574	standard; protein; 369 AA.
XX	ABG13574	
XX	ABG13574;	
DT	18-FEB-2002	(first entry)
XX		
DE	Novel human diagnostic protein #13565.	
XX		
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
KW	Food supplement; medical imaging; diagnostic; genetic disorder.	
XX		
OS	Homo sapiens.	
XX	MO200175067-A2.	
PN		
XX	11-OCT-2001.	
PD		
XX		
PF	30-MAR-2001; 2001WC-US008631.	
XX		
PR	31-MAR-2000; 2000US-00540217.	
XX	23-AUG-2000; 2000US-00649167.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Dymanac RT, Liu C, Tang YT,	
XX		
DR	WPI; 2001-639362/73.	
DR	N-PSDB; AAS77761.	
XX		
PT	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity.	
PS	Claim 20; SEQ ID NO 43933; 103bp; English.	
XX		
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)	
CC	sequences. (I) is useful as hybridisation probe, polymerase chain	
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	
CC	and in recombinant production of (II). The polynucleotides are also used	
CC	in diagnostics as expressed sequence tags for identifying expressed	
CC	genes. (I) is useful in gene therapy techniques to restore normal	

CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (II) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and other traits to assess biodiversity
 CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic
 CC patient did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIP0 at
 CC ftp.wipo.int/pub/published_pct_sequences

Sequence 369 AA;

Query Match 45.1%; Score 890; DB 4; Length 369;
 Best Local Similarity 48.6%; Pred. No. 8.9e-84;
 Matches 171; Conservative 75; Mismatches 100; Indels 6; Gaps 4;

6 KVFECAPNGKVTLYMGKRPVNDHSGVEPIDGIVLDEYIRDNKRVFGQIVCSFRRGR 65
 7 RVPFKASPNKGLTYLTKRDFVHDIDLPVDGVLTPEYLKE-RKVTYTLTCAFRYGR 65
 66 EEDVWGLNFOKEICLASEQIYPRPEKSDKEQTKLOEKLKKGSAIPTFTNISPNAPS 125
 66 EDDVGLTFRKDLFANVQSPFPAEDKPLTKQELTKRGEHAYPTFTFIPNLPIC 125
 126 SVTLQGGEDNDGPGCVSYVYKIFAGSETRTRHRSVTTLGIRKIQAPATKQGPCTL 125
 126 SVTLQGPEDTGKACGVDEAKFAENLEKIKHNSVGLVIRKQVAPEREGPFAE 185
 186 VRKDFMLSPGELLEVTLDKQLYLHGERIGVNICIRNNSKNKKIKKAMVQGVVFLQ 245
 186 TRQGLMSKPLHLEASLDKEIYYHGEPIGVNHNNTKIVETKIKISVRQYADICLEN 245
 246 NSGYRNTVASLETSEGCPIOPGSSLOKVMYLTPLSSNKRGGIALDQIKRQDCLAST 305
 246 TAQYKCPVAMEADD--TVAPSSIFCKVYTLTPFLANNEKRGALDQKLNHEDTNLAAS 303
 306 TLLAQPDORDAFGYIIISYAVYKFLFALGAGELSAELPFVLMHPKP-GTYA 355
 304 TLLHGANKEVILGIIIVSYKVKLVV-SRGGDVAVELPFTLHMFRAQGRTPA 354

RESULT 15

AAG67785 standard; protein; 382 AA.
 AC AAG67785;

DT 10-DEC-2001 (first entry)
 XX

DE An exemplary phosphorylation-independent arrestin mutant.

KM Arrestin; phosphorylation-independent arrestin mutant.
 KM G protein-coupled receptor; GPCR; GPCR ligand; retinitis pigmentosa;
 KM stationary night blindness; colour blindness; nephrogenic DI;
 KM familial hypocalcaemic hypocalcaemia; hyperparathyroidism;
 KM neurological disorder.

OS Unidentified.

XX WO200167106-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US007304.

XX 03-MAR-2000; 2000US-0186706P.

XX (MILL-) MILLENMUM PHARM INC.
 PA
 XX
 XX
 PT Berstein G;
 XX
 XX
 DR WPI; 2001-602637/68.

PT Identifying a G protein-coupled receptor ligand, useful for treating e.g.
 PT retinitis pigmentosa, color blindness or neurological disorders, uses
 PT phosphorylation-independent arrestin mutants particularly suited for in
 XX vitro screening assays.

PS Disclosure; Page 47; 47pp; English.

CC The present sequence represents an exemplary phosphorylation-independent
 CC arrestin mutant. Such mutants are used in screening assays to identify
 CC ligands and/or modulators of G protein-coupled receptors (GPCRs). A
 CC method for identifying a GPCR ligand comprises contacting a composition
 CC comprising the GPCR and a constitutively active arrestin mutant with a
 CC test compound; and determining the ability of the test compound to
 CC modulate binding of arrestin mutant to the GPCR, where modulation of
 CC binding indicates that the test compound is a GPCR ligand. The method is
 CC used for identifying potential ligands and/or modulators of GPCRs,
 CC particularly modulators, for use as human therapeutics. Modulators of
 CC GPCR may be used for treating patients having e.g. retinitis pigmentosa,
 CC stationary night blindness, colour blindness, nephrogenic DI, isolated
 CC hypocalcaemic hypocalcaemia, hyperparathyroidism and neurological
 CC disorders. The methods may be used for screening pluralities of test
 CC compounds (e.g. a small molecule library of compounds) or a composition
 CC containing a plurality of GPCRs. The methods may be used in screening
 CC assays for identification of natural and surrogate agonists of orphan
 CC GPCRs, and for identification of GPCR antagonists and/or agonists

Sequence 382 AA;

Query Match 44.9%; Score 885; DB 4; Length 382;
 Best Local Similarity 48.3%; Pred. No. 3.1e-83;
 Matches 167; Conservative 77; Mismatches 98; Indels 4; Gaps 3;

6 KVFECAPNGKVTLYMGKRPVNDHSGVEPIDGIVLDEYIRDNKRVFGQIVCSFRRGR 65
 8 RVPFKASPNKGLTYLTKRDFVHDIDLPVDGVLTPEYLKE-RKVTYTLTCAFRYGR 66
 66 EEDVWGLNFOKEICLASEQIYPRPEKSDKEQTKLOEKLKKGSAIPTFTNISPNAPS 125
 67 EDDVGLTFRKDLFANVQSPFPAEDKPLTKQELTKRGEHAYPTFTFIPNLPIC 126
 126 SVTLQGGEDNDGPGCVSYVYKIFAGSETRTRHRSVTTLGIRKIQAPATKQGPCTL 126
 126 SVTLQGPEDTGKACGVDEAKFAENLEKIKHNSVGLVIRKQVAPEREGPFAE 185
 127 SVTLQGPEDTGKACGVDEAKFAENLEKIKHNSVGLVIRKQVAPEREGPFAE 186
 186 VRKDFMLSPGELLEVTLDKQLYLHGERIGVNICIRNNSKNKKIKKAMVQGVVFLQ 245
 187 TTRHFLMSDSILHLEASLDKEIYYHGEPIGVNHNNTKIVETKIKISVRQYADICLEN 246
 246 NSGYRNTVASLETSEGCPIOPGSSLOKVMYLTPLSSNKRGGIALDQIKRQDCLAST 305
 247 TAQYKCPVAMEADD--TVAPSSIFCKVYTLTPFLANNEKRGALDQKLNHEDTNLAAS 304
 306 TLLAQPDORDAFGYIIISYAVYKFLFALGAGELSAELPFVLMHPKP 351
 305 TIVKGANKEVILGIIIVSYKVKLVV-SRGGDVAVELPFTLHMFRAQGRTPA 349

Search completed: February 9, 2005, 23:42:08
 Job time : 167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:44:27 ; Search time 43 Seconds
(without alignments)
664.897 Million cell updates/sec

Title: US-10-056-405-2

Perfect score: 1973

Sequence: 1 MYVNFVFKKCAPNGKVTLY.....ADVFTRQPTIDQASVDFE 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907.5	46.0	410	US-09-880-137-3	Sequence 3, Appl1
2	903.5	45.8	418	US-09-880-137-1	Sequence 1, Appl1
3	897.5	45.5	418	US-09-880-137-2	Sequence 2, Appl1
4	885	44.9	382	US-09-880-137-8	Sequence 8, Appl1
5	885	44.9	409	US-09-880-137-4	Sequence 4, Appl1
6	880	44.6	409	US-09-880-137-7	Sequence 7, Appl1
7	853.5	43.3	401	US-09-949-016-8383	Sequence 8383, Ap
8	853.5	43.3	401	US-09-949-016-8384	Sequence 8384, Ap
9	780.5	39.6	388	US-09-880-137-5	Sequence 6, Appl1
10	779.5	39.5	388	US-09-880-137-5	Sequence 5, Appl1
11	767.5	38.9	315	US-09-949-016-9940	Sequence 9940, Ap
12	767.5	38.9	315	US-09-949-016-9940	Sequence 9940, Ap
13	316	16.0	197	US-09-270-767-58832	Sequence 58832, A
14	225	11.4	137	US-09-270-767-32381	Sequence 32381, A
15	225	11.4	137	US-09-270-767-47598	Sequence 47598, A
16	101	5.1	363	US-09-270-767-44161	Sequence 44161, A
17	99.5	5.0	600	US-09-569-037-4	Sequence 4, Appl1
18	95.5	4.8	1034	US-09-543-681A-8172	Sequence 8172, Ap
19	92.5	4.7	1651	US-08-447-411-2	Sequence 2, Appl1
20	91.5	4.6	866	US-09-007-476-2	Sequence 10219, A
21	91	4.5	866	US-09-949-016-10219	Sequence 4363, Ap
22	89.5	4.5	539	US-09-134-000C-4363	Sequence 2, Appl1
23	89.5	4.5	725	US-08-425-843-2	Sequence 2, Appl1
24	88	4.4	1729	US-09-553-690-2	Sequence 775, App
25	86.5	4.4	684	US-09-636-215-775	Sequence 775, App
26	86.5	4.4	684	US-09-685-166A-775	Sequence 775, App
27	86.5	4.4	684	US-09-679-426-775	Sequence 775, App

28	86.5	4.4	684	US-09-759-143-775	Sequence 775, App
29	86.5	4.4	684	US-09-651-236-775	Sequence 775, App
30	86.5	4.4	1030	US-09-091-117-2	Sequence 2, Appl1
31	86	4.4	657	US-09-248-796A-19232	Sequence 19232, A
32	86	4.4	714	US-09-422-869-22	Sequence 22, Appl1
33	86	4.4	714	US-09-538-092-867	Sequence 867, App
34	85.5	4.3	800	US-09-583-110-2916	Sequence 2916, Ap
35	85.5	4.3	801	US-09-107-433-4536	Sequence 4536, Ap
36	85	4.3	477	US-09-489-039A-14257	Sequence 14257, A
37	85	4.3	783	US-08-843-521-2	Sequence 43701, A
38	85	4.3	783	US-09-012-871-2	Sequence 2, Appl1
39	85	4.3	1137	US-09-252-991A-24829	Sequence 24829, A
40	84.5	4.3	415	US-09-134-001C-4817	Sequence 4817, Ap
41	84.5	4.3	493	US-09-543-681A-7520	Sequence 7520, Ap
42	84.5	4.3	511	US-09-489-039A-14257	Sequence 14257, A
43	84.5	4.3	943	US-09-902-540-10641	Sequence 10641, A
44	84	4.3	626	US-09-248-796A-14855	Sequence 14855, A
45	84	4.3	1702	US-08-296-791-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1					
US-09-880-137-3					
; Sequence 3, Application US/09860137					
; Patent No. 6640025					
; GENERAL INFORMATION:					
; APPLICANT: Berstein, Gabriel					
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G					
; FILE REFERENCE: NMI-131					
; CURRENT APPLICATION NUMBER: US/09/880,137					
; PRIOR FILING DATE: 2001-03-05					
; PRIOR FILING DATE: 2000-03-03					
; NUMBER OF SEQ ID NOS: 8					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 3					
; LENGTH: 410					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-09-880-137-3					
Query Match					
Best Local Similarity 47.2%; Pred. No. 2.7e-87;					
Matches 176; Conservative 78; Mismatches 112; Indels 7; Gaps 4;					
Qy	6	KVFCKAPNGKVTLYMGKRDVDVSGVEPIDGIYVLDDEYIRDNRKVFQIVCSFRYGR	65		
Db	7	RVFCKASPNKCLTVYIGKRFVDHIDLVDPDGVLVDPDYLYKE-RRVYVTLTCAFRYGR	65		
Qy	66	EEDEVMGLNFQKELCLASEQIYPRPEKSKDEQTKLOERLLKKGSMNIPPTFNISPNAPS	125		
Db	66	EDLDVIGLTFKRLKLFVAVQSPFPAPEDEKPLRLRLRLKKGGEHAYPTFELPPMLPC	125		
Qy	126	SVTLQGEDNDGPCGVSYVYKIFAGESETRDHRBSYVTLTGRIKIOFATKQGGQCCCTL	185		
Db	126	SVTLQPEPTGTAGCGVDVFKAFCAENLEKHKRNSVALVIRKQYAPARERPPPTAE	185		
Qy	186	VRKDFMLSPGELEVTYLDKQLYHGERIGVNCIRNNSMKWYKIKAMVQGVVVLFQ	245		
Db	186	TTQGFMLSDKRLHLSLDEKIYYHGPISVNVHVTNTKTKIKISVRYADICLFN	245		
Qy	246	NGSIRNTVASLETSCEGPIPGSSLQKVMYLTPLLSNKKORGLADGQIKRQDCLAST	305		
Db	246	TAQYKCPVAMEADD--TVAPSSITFCVKVYTLTFPLANNRKGIALDGKIKHEDNTIASS	303		
Qy	306	TLAAGDQDPAFVITISYAVKVLPLGALGGEASLFPVLMHPKPGTK--AKVIHADS	362		
Db	304	TLIRGANREITLIISYKVKVL-VBSRGDVAVELPFTLMPKPKGEFPHHVEPNET	362		
Qy	363	QADVFTRQPTID 375			


```

Qy 126 SVTLQOGEDBDNDGDCGVSYVYKIKPAGESEPTDRTRSTVTLGIRKIQAPTRKGGQDPCTL 185
Db 127 SVTLQOPGEDDCKACGVDFEIRAFPAKSLSEKSHRNSVRLVIRVQFAPEKPGQPDSAE 186
Qy 186 VRKDFMSPGELSELEVTLDKQLYLHGERIGVNICIRNNSNKKVKKIKAMVOGVVYVLFQ 245
Db 187 TTRHFLMSDRSLHLEASLDKELYHGEPLANNVAVTNNSTKTVKKIKSVQYADICLFS 246
Qy 246 NGSYRNTVASLETSEGCPIQPGSSIQKVMYITPLLSNKKORGIALDQIKRQDCLAST 305
Db 247 TAOYKPCPAQLEQDD -QVSPSSTCKYTYITPLSDNRKRGALDQKLKHEDTNLASS 304
Qy 306 TLLAQPDQRDAFGVIISAVVYKFLGLAGELSALEPFVLMHPK 351
Db 305 TIYEGANKEVYLGILVSYRVKVLV -SRGDVSVLELPFVLMHPK 349

RESULT 5
US-09-880-137-4
; Sequence 4, Application US/09880137
; Patent No. 6640025
; GENERAL INFORMATION:
; APPLICANT: Barsteirn, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US/09/880.137
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-137-4

Query Match 44.9%; Score 885; DB 4; Length 409;
Best Local Similarity 48.3%; Pred. No. 6, 6e-85;
Matches 167; Conservative 77; Mismatches 98; Indels 4; Gaps 3;

Qy 6 KVFKKCAENGKVTLYMGRRDPVDVHGVPEPIDGIVLDDXYIRDNKRVFGQIVCSFRYGR 65
Db 8 RVFKKSPNCKLTIVYLGKRFDFVHDLDKVDPDVGVLVDPDYLKD -RKVFVTLTCAFRYGR 66
Qy 66 FEDRYMGINFOXELCLASEQIYPPREKSDKEQTKIQEQLKLKLSNAPLPTFNISPAPNS 125
Db 67 EDLDVLGSLSPKFDPIATYQAFPPPNPDRPTRLQDRLLRLCGHAPFFETIIPQNLPC 126
Qy 126 SVTLQOGEDBDNDGDCGVSYVYKIKPAGESEPTDRTRSTVTLGIRKIQAPTRKGGQDPCTL 185
Db 127 SVTLQOPGEDDCKACGVDFEIRAFPAKSLSEKSHRNSVRLVIRVQFAPEKPGQPDSAE 186
Qy 186 VRKDFMSPGELSELEVTLDKQLYLHGERIGVNICIRNNSNKKVKKIKAMVOGVVYVLFQ 245
Db 187 TTRHFLMSDRSLHLEASLDKELYHGEPLANNVAVTNNSTKTVKKIKSVQYADICLFS 246
Qy 246 NGSYRNTVASLETSEGCPIQPGSSIQKVMYITPLLSNKKORGIALDQIKRQDCLAST 305
Db 247 TAOYKPCPAQLEQDD -QVSPSSTCKYTYITPLSDNRKRGALDQKLKHEDTNLASS 304
Qy 306 TLLAQPDQRDAFGVIISAVVYKFLGLAGELSALEPFVLMHPK 351
Db 305 TIYEGANKEVYLGILVSYRVKVLV -SRGDVSVLELPFVLMHPK 349

RESULT 6
US-09-880-137-7
; Sequence 7, Application US/09880137
; Patent No. 6640025
; GENERAL INFORMATION:

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Best Local Similarity 46.1%; Pred. No. 1,4e-81;
Matches 170; Conservative 75; Mismatches 111; Indels 13; Gaps 5;

```

QY 17 VTLVGRDVPDHDVSGVEPIDGIYVLDDEYIRDNKRVFGQIVCSFRYGRGEEDEWVGNLQ 76
D 1 LTVVLRGRDVPDHDVLDVDPVGVVLDPEYLKE--RRVYVTLTCARFYGRGEEDELDVGLTFR 59
QY 77 KELCLASEQIYPRPEKSDKEQOTLQERLLKLGSNAPPTFNISNMASSVTLQOGEEDN 136
D 60 KDLFVANVOSFPAPAPEDKPLTRLOERLLKLGSEHAFFTEIPNLCSVTLQPGEDT 119
QY 137 GDCGVSYYVKIPAGESETDRHRSSTVTLGIRKIQAPATKQGOOPCTVLAKDPMLSPE 196
D 120 GKACGVDEYKAFCAENLEEKIKHNSVRLVIRKQVAPERRPGPPTAETTRQGLMSDKP 179
QY 197 LHEVTLDKQYLHGERIGVNICIRNNSKNKWKIKAMVQGVVULFONGSTRNTYASL 256
D 180 LHEASLDKEIYYHGEPISSVNVHTNNKTKIKISVROYADICLFNTAQKCPVAME 239
QY 257 ETSBGPPIQSSLOKMWYLPILSSNKGORGIALDQIKRODOCLASTTLAQPOQORDA 316
D 240 EADD--TVAPSTFCKYVTLTPFLANNREKRGALDGLKHEPTNLASTLLREGANREI 297
QY 317 FGVISYAVKVLFL--GALGELS-----ELPFVLMHPRPGTK--AKVHADSQADV 366
D 298 LGIIVSYKVKYLVVSRGGLGLDGLASSDVAVELPFTLMHPKPEBPBREVENETPVDT 357
QY 367 ETRFODTID 375
D 358 NLIELDTND 366

```

RESULT 8

US-09-949-016-8384
Sequence 8384, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8384
LENGTH: 401
TYPE: PRT
ORGANISM: Human
US-09-949-016-8384

Query Match 43.3%; Score 853.5; DB 4; Length 401;
Best Local Similarity 46.1%; Pred. No. 1.4e-81;
Matches 170; Conservative 75; Mismatches 111; Indels 13; Gaps 5;

```

QY 17 VTLVGRDVPDHDVSGVEPIDGIYVLDDEYIRDNKRVFGQIVCSFRYGRGEEDEWVGNLQ 76
D 1 LTVVLRGRDVPDHDVLDVDPVGVVLDPEYLKE--RRVYVTLTCARFYGRGEEDELDVGLTFR 59
QY 77 KELCLASEQIYPRPEKSDKEQOTLQERLLKLGSNAPPTFNISNMASSVTLQOGEEDN 136
D 60 KDLFVANVOSFPAPAPEDKPLTRLOERLLKLGSEHAFFTEIPNLCSVTLQPGEDT 119
QY 137 GDCGVSYYVKIPAGESETDRHRSSTVTLGIRKIQAPATKQGOOPCTVLAKDPMLSPE 196
D 120 GKACGVDEYKAFCAENLEEKIKHNSVRLVIRKQVAPERRPGPPTAETTRQGLMSDKP 179

```

```

QY 197 LHEVTLDKQYLHGERIGVNICIRNNSKNKWKIKAMVQGVVULFONGSTRNTYASL 256
D 180 LHEASLDKEIYYHGEPISSVNVHTNNKTKIKISVROYADICLFNTAQKCPVAME 239
QY 257 ETSBGPPIQSSLOKMWYLPILSSNKGORGIALDQIKRODOCLASTTLAQPOQORDA 316
D 240 EADD--TVAPSTFCKYVTLTPFLANNREKRGALDGLKHEPTNLASTLLREGANREI 297
QY 317 FGVISYAVKVLFL--GALGELS-----ELPFVLMHPRPGTK--AKVHADSQADV 366
D 298 LGIIVSYKVKYLVVSRGGLGLDGLASSDVAVELPFTLMHPKPEBPBREVENETPVDT 357
QY 367 ETRFODTID 375
D 358 NLIELDTND 366

```

RESULT 9

US-09-880-137-6
Sequence 6, Application US/09880137
Patent No. 6640025
GENERAL INFORMATION:
APPLICANT: Bernstein, Gabriel
TITLE OF INVENTION: METHODS OF ASSAYING FOR G
FILE REFERENCE: NMI-131
CURRENT APPLICATION NUMBER: US/09/880,137
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/186,706
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 388
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-137-6

Query Match 39.6%; Score 780.5; DB 4; Length 388;
Best Local Similarity 42.7%; Pred. No. 7.3e-74;
Matches 166; Conservative 79; Mismatches 125; Indels 19; Gaps 7;

```

QY 6 KVFKKCAPNGKVTLYMKRDPVDHVSVEPIDGIYVLDDEYIRDNKRVFGQIVCSFRYGR 65
D 3 KVFKKTSNGKLSIYKGDVFDVDTVPEDGVVLDPEYLK--CRKLFWMLTCAFRYGR 61
QY 66 EEDVWGNLQERLLKLGSEHAFFTEIPNLCSVTLQPGEDT 124
D 62 DDEEVGLTFRKQYVOTLQVVPASSSSPQALTVLQERLLKLGSEHAFFTEIPNLCSVTLQPGEDT 121
QY 125 SSVTLQOGEEDNPGCVSYYVKIPAGESETDRHRSSTVTLGIRKIQAPATKQGOOPCT 184
D 122 CSVTLQPGEDAGKCGIDFVYKSCAENPEYTSKDYVRLVVRKQVAPERRPGPPTAETTRQGLMSDKP 181
QY 185 LVRKQDPMLSPELEVTLDKQYLHGERIGVNICIRNNSKNKWKIKAMVQGVVULFONGSTRNTYASL 244
D 182 QTRIRFLLSAQPLQIQAAMMDREYVHGEPISSVNVHTNNKTKIKISVROYADICLFNTAQKCPVAME 239
QY 245 QNGSYRNTYASLETSBGPPIQSSLOKMWYLPILSSNKGORGIALDQIKRODOCLAS 304
D 242 SLDKTKTVFIOEFT--TVANSSFSQSAFVPTTLAASQKRGALDGLKHEPTNLASTLLREGANREI 299
QY 305 TLLAQPOQORDAFGVISYAVKVLFL--GALGELS-----ELPFVLMHPRPGTKAKV 357
D 300 STIRIRPGMDKELGIIIVSYKVKYLVVSRGGLGLDGLASSDVAVELPFTLMHPKPEBPBREVENETPVDT 357
QY 358 IHADSQADV--ETRFODTIDQOASVDFE 383
D 359 ---SSEDIYIEBTRKGESESQAVAE 383

```

RESULT 10

US-09-880-137-5


```

: Sequence 5, Application US/09880137
: Patent No. 6640025
: GENERAL INFORMATION:
: APPLICANT: Berstein, Gabriel
: TITLE OF INVENTION: METHODS OF ASSAYING FOR G
: TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
: FILE REFERENCE: NMI-131
: CURRENT APPLICATION NUMBER: US/09/880.137
: CURRENT FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: US 60/186,706
: PRIOR FILING DATE: 2000-03-03
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PasteSeq for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 388
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-880-137-5

```

Query Match	39.5%	Score 779.5;	DB 4;	Length 388;
Boot Local Similarity	42.4%;	Pred. NO. 9.3e-74;		
Matches 165;	Conservative 80;	Mismatches 125;	Indels 19;	Gaps 7;

[illegible]

```

RESULT 11
US-09-949-016-9940
; Sequence 9940, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO0107
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: ParSeq for Windows Version 4.0
; SEQ ID NO 9940
; LENGTH: 315

```

```

; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9940

```

Query Match	38.9%	Score 767.5;	DB 4;	Length 315;
Best Local Similarity	48.0%	Pred. No. 1.2e-72;		
Matches 145; Conservative	66;	Mismatches 88;	Indels 3;	Gaps 2;

Qy	6	KXFKKCA	PNB	GKXTL	LMGR	BD	PVDH	NSG	VEP	LDG	IVL	DDEX	IPBN	RV	FGQ	IVCS	PFYGR	65																																						
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																																							
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																																							
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																																							
Db	8	R	VFKFS	CPCK	LT	VY	LGRD	FVDH	DKD	BD	PVDG	VVL	VD	PD	LLK	-R	V	FTL	TC	AEPYGR	66																																			
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:																																				
Qy	66	E	EED	VW	GLN	FQ	KEL	CLASE	QI	Y	P	R	P	R	E	K	S	D	K	O	T	L	O	E	R	L	L	K	L	G	S	N	A	I	P	E	T	F	N	I	S	P	A	P	S	1255										
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:													
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:													
Db	67	E	D	D	V	I	G	L	S	F	R	D	L	T	A	T	Y	O	A	F	R	P	P	N	P	R	P	T	L	O	B	R	L	K	L	G	O	N	H	P	F	F	T	I	P	O	V	L	P	C	126					
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:										
Qy	126	S	V	T	L	O	O	G	E	D	N	D	P	C	G	S	V	Y	V	K	I	P	A	G	E	S	E	T	R	T	R	S	T	V	T	L	G	I	R	K	I	O	P	A	R	T	O	O	O	P	C	T	L	185		
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:							
Db	127	S	V	T	L	O	G	E	P	E	D	T	O	K	A	G	V	D	P	E	I	R	A	F	C	A	N	S	L	E	K	S	H	K	N	S	V	R	L	I	R	K	V	P	A	R	E	K	O	P	S	A	E	186		
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:							
Qy	186	V	R	K	D	E	M	L	S	F	G	E	L	E	V	T	L	D	K	O	L	Y	L	H	E	R	I	G	N	I	C	I	R	N	N	S	N	K	M	V	K	I	T	A	M	Y	O	O	G	V	D	V	L	F	Q	245
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:							
Db	187	T	R	H	F	M	S	D	R	S	L	H	E	A	S	L	D	K	E	L	Y	H	G	E	P	L	N	N	V	H	T	N	N	S	T	K	V	Y	K	I	K	V	S	R	O	A	D	I	C	L	F	S	246			
Qy	246	N	S	G	I	R	T	M	V	A	S	L	E	T	S	G	C	P	I	O	R	E	S	S	I	O	K	M	T	L	P	L	L	S	N	K	O	R	G	A	L	D	G	O	I	K	R	O	D	C	L	A	S	T	305	
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:							
Db	247	T	A	O	K	C	P	V	A	O	L	E	O	D	-	O	V	S	P																																					

```

RESULT 12
US-09-270-767-58832
; Sequence 58832, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58832
; LENGTH: 197
; TYPE: PR1
; ORGANISM: Drosophila melanogaster
; US-09-270-767-58832

```

```

Query Match      16.0% Score 316; DB 4; Length 197;
Best Local Similarity 49.6%; Pred. No. 4,3e-25;
Matches 55; Conservative 25; Mismatches 33; Indels 2; Gaps 1;

QY      235 VQGVGVVLFQNGSYRNTVAFLSETECCPIQPGSSLOKMYLTPLESSNKQRGRIALDQ 294
          :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      2 VRRGADICLFSTAQYKSKVAIEISEDCQYVAPFTLSKVFELCPULANNKDKWGLDQ 61

QY      295 IKRQDQCLASTTLTAQPDORDAFVITISYAVKVKLFLGA--LGGELSNELPFVLMHPP 351
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      62 LKHEDTLASTLITNPAQRESLGIWVHKYKVKLLSSPLNLGDLVAELPFTLMHPP 120

RESULT 13
US-09-270-767-43475
; Sequence 43475, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

```

```

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43475
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43475

```

```

Query Match
Best Local Similarity 16.0%; Score 316; DB 4; Length 475;
Matches 59; Conservative 25; Mismatches 33; Indels 2; Gaps 1;

```

```

QY 235 VQGVVULFQNGSYNNVVASLSESGCPIQSGSLQKWTLPPLLSNKKORGIADQ 294
DB 280 VRFPADICLFTSAQKSVVAIESEDCQVAPGFTLSKVEFLCPLANNKDKWGLADQ 339
QY 295 IRRQDCLASTTLAQPORDAFGVIISYAVKYLFLGA--LGGELSAELPVLNHPKP 351
DB 340 LKHEDTNLASSTLITNPAQRSLGIMVHKVKYKLLISSPLNGDLVLELFTLMHPKP 398

```

```

RESULT 14
US-09-270-767-32381

```

```

; Sequence 32381, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

```

```

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32381
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32381

```

```

Query Match
Best Local Similarity 11.4%; Score 225; DB 4; Length 137;
Matches 52; Conservative 26; Mismatches 38; Indels 6; Gaps 4;

```

```

QY 6 KVFKKCAPNGKVTLYMGKRDVVDHVGVEP--IDGIVLDDERYRDNKRVFGQIVCSFRY 63
DB 19 RYVKKTSPPCVLTLYLPTREIT--LTGNNSVLRGIVYVDPKAIQGR--VYAOULTLTFRY 75
QY 64 GREDEVMGLNFOKEICLASEQIYPRPEKSDKEO--TKLOERLLKLGSNALPFTFNISPN 122
DB 76 GREDEVMGLNFCNEAIMSLHQIWPRLBEPPELSPLQEALMKRLGDGAHPFTLSLSY 135
QY 123 AP 124
DB 136 AP 137

```

```

RESULT 15
US-09-270-767-47598

```

```

; Sequence 47598, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

```

```

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47598
; LENGTH: 137

```

```

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47598

```

```

Query Match
Best Local Similarity 11.4%; Score 225; DB 4; Length 137;
Matches 52; Conservative 26; Mismatches 38; Indels 6; Gaps 4;

```

```

QY 6 KVFKKCAPNGKVTLYMGKRDVVDHVGVEP--IDGIVLDDERYRDNKRVFGQIVCSFRY 63
DB 19 RYVKKTSPPCVLTLYLPTREIT--LTGNNSVLRGIVYVDPKAIQGR--VYAOULTLTFRY 75
QY 64 GREDEVMGLNFOKEICLASEQIYPRPEKSDKEO--TKLOERLLKLGSNALPFTFNISPN 122
DB 76 GREDEVMGLNFCNEAIMSLHQIWPRLBEPPELSPLQEALMKRLGDGAHPFTLSLSY 135
QY 123 AP 124
DB 136 AP 137

```

```

Search completed: February 9, 2005, 23:55:39
Job time : 45 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 9, 2005, 23:23:01 ; Search time 130 Seconds
(without alignments)
959.569 Million cell updates/sec

Title: US-10-056-405-2

Perfect score: 1973

Sequence: 1 MYVNFVKFKKCAENGKVTLY.....ADVFTRQDTIDQASVDFE 383

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1973	100.0	383	14	US-10-094-240-2
2	1973	100.0	383	14	US-10-056-405-2
3	1023	51.8	389	14	US-10-094-240-25
4	982.5	49.8	401	14	US-10-094-240-27
5	907.5	46.0	410	9	US-09-880-137-3
6	907.5	46.0	410	10	US-09-800-137A-3
7	903.5	45.8	418	9	US-09-880-137-1
8	903.5	45.8	418	10	US-09-800-137A-1
9	903.5	45.8	452	14	US-10-038-010-52
10	897.5	45.5	418	14	US-10-043-487-348
11	897.5	45.5	418	9	US-09-880-137-2
12	897.5	45.5	418	10	US-09-800-137A-2
13	885	44.9	382	9	US-09-880-137-8

14	885	44.9	382	10	US-09-800-137A-8	Sequence 8, Appl1
15	885	44.9	409	9	US-09-880-137-4	Sequence 4, Appl1
16	885	44.9	409	9	US-09-800-137A-4	Sequence 4, Appl1
17	880	44.6	409	9	US-09-880-137-7	Sequence 7, Appl1
18	880	44.6	409	10	US-09-800-137A-7	Sequence 7, Appl1
19	869.5	44.1	410	14	US-10-038-010-54	Sequence 54, Appl
20	865	43.8	454	15	US-10-369-493-6702	Sequence 6702, Ap
21	780.5	39.6	388	9	US-09-880-137-6	Sequence 6, Appl1
22	780.5	39.6	388	10	US-09-800-137A-6	Sequence 6, Appl1
23	779.5	39.5	388	9	US-09-880-137-5	Sequence 5, Appl1
24	779.5	39.5	388	10	US-09-800-137A-5	Sequence 5, Appl1
25	379.5	19.2	466	16	US-10-722-357-31	Sequence 31, Appl
26	361.5	18.3	212	14	US-10-106-698-4673	Sequence 4673, Ap
27	173	8.8	65	9	US-09-864-761-45375	Sequence 45375, A
28	111	5.6	38	9	US-09-864-761-40768	Sequence 40768, A
29	98.5	5.0	349	15	US-10-335-977-5814	Sequence 5814, Ap
30	98.5	5.0	1632	16	US-10-437-963-177962	Sequence 177962,
31	95	4.8	348	15	US-10-424-599-201640	Sequence 201640,
32	94.5	4.8	1211	15	US-10-282-122A-69274	Sequence 69274, A
33	92	4.7	320	15	US-10-335-977-5813	Sequence 5813, Ap
34	92	4.7	944	15	US-10-282-122A-71708	Sequence 71708, A
35	92	4.7	1348	15	US-10-282-122A-56877	Sequence 56877, A
36	91.5	4.6	239	15	US-10-282-122A-51797	Sequence 51797, A
37	91.5	4.6	416	15	US-10-282-122A-44280	Sequence 44280, A
38	91.5	4.6	752	16	US-10-437-963-115259	Sequence 115259,
39	91.5	4.6	863	9	US-09-815-242-10623	Sequence 10623, A
40	91	4.6	488	15	US-10-369-493-20203	Sequence 20203, Ap
41	91	4.6	800	5	US-10-104-047-3534	Sequence 3534, Ap
42	90.5	4.6	422	9	US-09-881-752A-142	Sequence 142, App
43	90.5	4.6	492	15	US-10-282-122A-58846	Sequence 58846, A
44	90.5	4.6	976	16	US-10-437-963-201424	Sequence 201424,
45	90	4.6	356	14	US-10-153-398-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-094-240-2
; Sequence 2, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N8289
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US/10/094,240
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-094-240-2

Query Match 100.0%; Score 1973; DB 14; Length 383;
Best Local Similarity 100.0%; Pred. No. 4, 4e-176;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYVNFVKFKKCAENGKVTLYMGKRDYVDHYSGVPEIDGIYVLDDEYTRDRKRYFGQIVCS 60
DB 1 MYVNFVKFKKCAENGKVTLYMGKRDYVDHYSGVPEIDGIYVLDDEYTRDRKRYFGQIVCS 60
QY 61 FRYGREDEVMGNFQKELCLASEQIYPRPEKSDKQTLQERLLKKLGSGNAIPFTFNIS 120
DB 61 FRYGREDEVMGNFQKELCLASEQIYPRPEKSDKQTLQERLLKKLGSGNAIPFTFNIS 120
QY 121 PNASVTLQGGEDNDGDPGVSYYVKKIFAGESETDRTHRSVTYLGIRKIQFAPTKQGO 180

```

Db      121  PNAPSSVTLLQGGEDNDGDCGVSYYVKIFAGESETDTRHRSVTLLGIRKIQAPATKQO 180
Qy      181  OPTLVKRDPMLSPELELVTLDKOLYHGERIGVNICIRNNSNKKVKKIKAMVOQGV 240
Db      181  OPTLVKRDPMLSPELELVTLDKOLYHGERIGVNICIRNNSNKKVKKIKAMVOQGV 240
Qy      241  VVLFONGSYRNTVASLETSEGCPIQPGSSLOKWMYLPPLSSNKKORGIALDGOIKRODQ 300
Db      241  VVLFONGSYRNTVASLETSEGCPIQPGSSLOKWMYLPPLSSNKKORGIALDGOIKRODQ 300
Qy      301  CLASTTLLAOPDORDAFGVIISYAVVKLFLGALGELSLELPVLMHPKGTAKVIHA 360
Db      301  CLASTTLLAOPDORDAFGVIISYAVVKLFLGALGELSLELPVLMHPKGTAKVIHA 360
Qy      361  DSQADVETFRDPTIDQASVDPE 383
Db      361  DSQADVETFRDPTIDQASVDPE 383

```

```

RESULT 2
US-10-056-405-2
/ Sequence 2, Application US/10056405
/ Publication No. US20030166013A1
/ GENERAL INFORMATION:
/ APPLICANT: ZWIEBEL, LAURENCE J.
/ TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
/ FILE REFERENCE: N7841
/ CURRENT APPLICATION NUMBER: US/10/056,405
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: 60/264,649
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 383
/ TYPE: PRT
/ ORGANISM: Anopheles gambiae
US-10-056-405-2

```

```

Query Match
Best Local Similarity 100.0%; Score 1973; DB 14; Length 383;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  MYNFKVFKKCAPNGKVTLLYMGKDFVDHSGVEPIDIGIVLDDXYTRDNKRVFGQIVCS 60
Db      1  MYNFKVFKKCAPNGKVTLLYMGKDFVDHSGVEPIDIGIVLDDXYTRDNKRVFGQIVCS 60
Qy      61  FRYGREDEVMGLNFQKELCLASEQIYRPERKSDKEQTKLQERLLKLGSAIIPFTNIS 120
Db      61  FRYGREDEVMGLNFQKELCLASEQIYRPERKSDKEQTKLQERLLKLGSAIIPFTNIS 120
Qy      121  PNAPSSVTLLQGGEDNDGDCGVSYYVKIFAGESETDTRHRSVTLLGIRKIQAPATKQO 180
Db      121  PNAPSSVTLLQGGEDNDGDCGVSYYVKIFAGESETDTRHRSVTLLGIRKIQAPATKQO 180
Qy      181  OPTLVKRDPMLSPELELVTLDKOLYHGERIGVNICIRNNSNKKVKKIKAMVOQGV 240
Db      181  OPTLVKRDPMLSPELELVTLDKOLYHGERIGVNICIRNNSNKKVKKIKAMVOQGV 240
Qy      241  VVLFONGSYRNTVASLETSEGCPIQPGSSLOKWMYLPPLSSNKKORGIALDGOIKRODQ 300
Db      241  VVLFONGSYRNTVASLETSEGCPIQPGSSLOKWMYLPPLSSNKKORGIALDGOIKRODQ 300
Qy      301  CLASTTLLAOPDORDAFGVIISYAVVKLFLGALGELSLELPVLMHPKGTAKVIHA 360
Db      301  CLASTTLLAOPDORDAFGVIISYAVVKLFLGALGELSLELPVLMHPKGTAKVIHA 360
Qy      361  DSQADVETFRDPTIDQASVDPE 383
Db      361  DSQADVETFRDPTIDQASVDPE 383

```

```

RESULT 3
US-10-094-240-25
/ Sequence 25, Application US/10094240
/ Publication No. US20030082637A1
/ GENERAL INFORMATION:
/ APPLICANT: ZWIEBEL, LAURENCE J.
/ TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
/ FILE REFERENCE: N8289
/ CURRENT APPLICATION NUMBER: US/10/094,240
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 10/056,405
/ PRIOR FILING DATE: 2002-01-24
/ NUMBER OF SEQ ID NOS: 60/264,649
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 25
/ LENGTH: 398
/ TYPE: PRT
/ ORGANISM: Anopheles gambiae
US-10-094-240-25

```

```

Query Match
Best Local Similarity 51.8%; Score 1023; DB 14; Length 398;
Matches 204; Conservative 65; Mismatches 111; Indels 10; Gaps 7;
Qy      1  MYNFKVFKKCAPNGKVTLLYMGKDFVDHSGVEPIDIGIVLDDXYTRDNKRVFGQIVCS 60
Db      1  MYNFKVFKKCAPNGKVTLLYMGKDFVDHSGVEPIDIGIVLDDXYTRDNKRVFGQIVCS 60
Qy      61  FRYGREDEVMGLNFQKELCLASEQIYRPERKSDKEQTKLQERLLKLGSAIIPFTNIS 120
Db      61  FRYGREDEVMGLNFQKELCLASEQIYRPERKSDKEQTKLQERLLKLGSAIIPFTNIS 120
Qy      121  PNAPSSVTLLQGGEDNDGDCGVSYYVKIFAGESETDTRHRSVTLLGIRKIQAPATKQO 180
Db      121  PNAPSSVTLLQGGEDNDGDCGVSYYVKIFAGESETDTRHRSVTLLGIRKIQAPATKQO 180
Qy      181  OPTLVKRDPMLSPELELVTLDKOLYHGERIGVNICIRNNSNKKVKKIKAMVOQGV 240
Db      181  OPTLVKRDPMLSPELELVTLDKOLYHGERIGVNICIRNNSNKKVKKIKAMVOQGV 240
Qy      241  VVLFONGSYRNTVASLETSEGCPIQPGSSLOKWMYLPPLSSNKKORGIALDGOIKRODQ 300
Db      241  VVLFONGSYRNTVASLETSEGCPIQPGSSLOKWMYLPPLSSNKKORGIALDGOIKRODQ 300
Qy      301  CLASTTLLAOPDORDAFGVIISYAVVKLFLGALGELSLELPVLMHPKGTAKVIHA 360
Db      301  CLASTTLLAOPDORDAFGVIISYAVVKLFLGALGELSLELPVLMHPKGTAKVIHA 360
Qy      361  DSQADVETFRDPTIDQASVDPE 383
Db      361  DSQADVETFRDPTIDQASVDPE 383

```

```

RESULT 4
US-10-094-240-27
/ Sequence 27, Application US/10094240
/ Publication No. US20030082637A1
/ GENERAL INFORMATION:
/ APPLICANT: ZWIEBEL, LAURENCE J.
/ TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
/ FILE REFERENCE: N8289
/ CURRENT APPLICATION NUMBER: US/10/094,240
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 10/056,405
/ PRIOR FILING DATE: 2002-01-24
/ NUMBER OF SEQ ID NOS: 60/264,649
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 27
/ LENGTH 401

```

```

; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-094-240-27

Query Match      49.8%; Score 982.5; DB 14; Length 401;
Best Local Similarity 46.7%; Pred. No. 3.3e-83;
Matches 191; Conservative 73; Mismatches 115; Indels 13; Gaps 6;

Qy 1 MVTYFKPKKCAKPGKATLYWYKGRDPVDHVSQVEPIGIVLDDDEYIRNRYKVFQIVCS 60
Db 1 MVTYVKPKKCAKPGKATLYWYKGRDPVDHVSQVEPIGIVLDDDEYIRNRYKVFQIVCS 59
Qy 61 FRYGRDEDEWGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKLGSAIPTFNIS 120
Db 60 YRYGRDEDEWGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKLGSAIPTFNIS 118
Qy 121 PNAASSVTLLQGGEDNDGPCVSYVYKIFAGESETDRTHRSSTVTLGIRKIQEPTQOQ 180
Db 119 PNPSSVTLLQGGEDNDGPCVSYVYKIFAGESETDRTHRSSTVTLGIRKIQEPTQOQ 178
Qy 181 Q-PCTLVKDDFMLSPELELEVTLDKOLYLHGERIGVNICIRNNSNMVKKIKAMVQGV 239
Db 179 RLPSLVKSGFTFSNGKISLEVTLDREIYHGEKTAATVOVSNKSKSVSIKCFIVQHT 238
Qy 240 DVVLFQNGSYRNTVASLETSEGCPIOPGSSLOKVMYTLPLSSNKKORGIALDQIQRD 299
Db 239 EITMV-NAQFSKIVAQLETKGECITTPGANLTKTFYLIPLAANKKHGIALDGHLEDED 297
Qy 300 QCLASTTLLAO-PQORDAFVYIISYAVKVLFLGALGELSABLPLVLMHPKPGT----- 353
Db 298 VNLASSTWVGSKSTGDAAGIVISYVAIKLNCGLTGEMQTDVFPKLLDPAPGTIEKR 357
Qy 354 ---KAVIHADSDQADVFETPRDQTDQASVDF 382
Db 358 SNAMKKMKSIQEHNRKVIYDDDDNIVFEDF 389

RESULT 5
US-09-880-137-3
; Sequence 3, Application US/09880137
; Patent No. US20020031295A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: NMI-131
; CURRENT APPLICATION NUMBER: US/09/880.137
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-137-3

Query Match      46.0%; Score 907.5; DB 9; Length 410;
Best Local Similarity 47.2%; Pred. No. 3.7e-76;
Matches 176; Conservative 78; Mismatches 112; Indels 7; Gaps 4;

Qy 6 KVFKKCAPNGKVTLYMGKRDVPHVSGVEPIDGIVLDDDEYIRNRYKVFQIVCSFRYGR 65
Db 7 RVFKKASPNKGLTYLTKRDPVDHIDLVDVPGVAVLDPETLKE-RRVYVTLTCAFRYGR 65
Qy 66 EEDVWGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKLGSAIPTFNISPNAPS 125
Db 66 EDDLVLGLTRKDLFVANVQSPFPAPEDKKPLTRLQERLLKLGSAIPTFNISPNAPS 125
Qy 126 SVTLQGGEDNDGPCVSYVYKIFAGESETDRTHRSSTVTLGIRKIQEPTQOQOQCTL 185
Db 126 SVTLQGGEDNDGPCVSYVYKIFAGESETDRTHRSSTVTLGIRKIQEPTQOQOQCTL 185

RESULT 6
US-09-800-137A-3
; Sequence 3, Application US/09800137A
; Publication No. US20030157553A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: NMI-131
; CURRENT APPLICATION NUMBER: US/09/800.137A
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-137A-3

Query Match      46.0%; Score 907.5; DB 10; Length 410;
Best Local Similarity 47.2%; Pred. No. 3.7e-76;
Matches 176; Conservative 78; Mismatches 112; Indels 7; Gaps 4;

Qy 6 KVFKKCAPNGKVTLYMGKRDVPHVSGVEPIDGIVLDDDEYIRNRYKVFQIVCSFRYGR 65
Db 7 RVFKKASPNKGLTYLTKRDPVDHIDLVDVPGVAVLDPETLKE-RRVYVTLTCAFRYGR 65
Qy 66 EEDVWGLNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKLGSAIPTFNISPNAPS 125
Db 66 EDDLVLGLTRKDLFVANVQSPFPAPEDKKPLTRLQERLLKLGSAIPTFNISPNAPS 125
Qy 126 SVTLQGGEDNDGPCVSYVYKIFAGESETDRTHRSSTVTLGIRKIQEPTQOQOQCTL 185
Db 126 SVTLQGGEDNDGPCVSYVYKIFAGESETDRTHRSSTVTLGIRKIQEPTQOQOQCTL 185
Qy 186 VRKDFMLSPELELEVTLDKOLYLHGERIGVNICIRNNSNMVKKIKAMVQGVVVLFPQ 245
Db 186 TTRQFLMSDKRPLHESALDEIYYHGEPISVNVAHVNTNTKTKYKKIKISRQVADICLFN 245
Qy 246 NSGYRNTVASLETSEGCPIOPGSSLOKVMYTLPLSSNKKORGIALDQIQRDQCLAST 305
Db 246 TAQYKCPVAMEEADD--TVAPSSTFCKVYTLTPFLANNREKRGLADGKLKHEDTNLASS 303
Qy 306 TLLAQPRQRAFGVYIISYAVKVLFLGALGELSABLPLVLMHPKPGTK--AKVIHADS 362
Db 304 TLLREGANREILIGIVSYKVKVL-VESRGSDVAVELPFTLMHPKPEBPBREVPENET 362
Qy 363 QADVFETPRDQTD 375
Db 363 PVDTNLIEDTND 375

RESULT 7
```

US-09-880-137-1
 ; Sequence 1, Application US/09880137
 ; Patent No. US20020031295A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berstein, Gabriel
 ; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
 ; FILE REFERENCE: NMT-131
 ; CURRENT APPLICATION NUMBER: US/09/880,137
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1: FASTSEQ for Windows Version 4.0
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Bos taurus
 ; US-09-880-137-1

Query Match
 Best Local Similarity 45.8%; Score 903.5; DB 9; Length 418;
 Matches 179; Conservative 76; Mismatches 112; Indels 13; Gaps 5;

QY 6 KYFKKCAPNGKVTLYMGKRDVDFVHSGVPEPIDGIVLDDERYIRDNKRVFGQIVCSFRYGR 65
 DB 7 RVFKKASPNKGLTVYGRDFVDHIDLVBPVGVVLDPEYLKE -RRVYVLTCAFRYGR 65
 QY 66 EDEVMGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKGSNAIPFTFNISPNAPS 125
 DB 66 EDLDVGLTFRKDLFVAVNQSFPAPEDKKPLTRLOERLLKGSNAIPFTFNISPNAPS 125
 QY 126 SVTLQGEDNDGDCGVSYYVYKIFAGESETDRTHRSVTYTGIRKIQAPFTKGOQOQPC 125
 DB 126 SVTLQGEDNDGDCGVSYYVYKIFAGESETDRTHRSVTYTGIRKIQAPFTKGOQOQPC 125
 QY 186 VKRDMFLSPGELLEVTLKQVLYHGERIGVNICIRNNSNMVYKIKAMVQGVVYLFQ 245
 DB 186 TTRQFLMSDPKHLHLSLDEKITYHGEPISVNHNVTNNKTKYKIKISVQYADICLFN 245
 QY 246 NGSYRNTVASLETSECPYIOPGSSLOKMYLTPLLSNKRGRGIALDQIKRQOCLAST 245
 DB 246 TAQKCPVAMEAD--TVAPSSYFCVYTLTFLANNREKGLALDQIKRQOCLAST 245
 QY 306 TLLAOPDQDAFGVYISYAVKVLFL--GALGELISA-----ELPVLAMPKRGTK--A 355
 DB 304 TLLRSGANREILGITYSVKVLVVSRGQLGLDGLASSDAVELPFTLMHPKKEPPHR 363
 QY 356 KYIHADSOADVETFRQDTID 375
 DB 364 EVPEHETPVDTMLIBLDTND 383

RESULT 8
 ; US-09-800-137A-1
 ; Sequence 1, Application US/09800137A
 ; Patent No. US2003015753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berstein, Gabriel
 ; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
 ; FILE REFERENCE: NMT-131
 ; CURRENT APPLICATION NUMBER: US/09/800,137A
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1: FASTSEQ for Windows Version 4.0
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Bos taurus
 ; US-09-800-137A-1

Query Match
 Best Local Similarity 45.8%; Score 903.5; DB 10; Length 418;
 Matches 179; Conservative 76; Mismatches 112; Indels 13; Gaps 5;

QY 6 KYFKKCAPNGKVTLYMGKRDVDFVHSGVPEPIDGIVLDDERYIRDNKRVFGQIVCSFRYGR 65
 DB 7 RVFKKASPNKGLTVYGRDFVDHIDLVBPVGVVLDPEYLKE -RRVYVLTCAFRYGR 65
 QY 66 EDEVMGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKGSNAIPFTFNISPNAPS 125
 DB 66 EDLDVGLTFRKDLFVAVNQSFPAPEDKKPLTRLOERLLKGSNAIPFTFNISPNAPS 125
 QY 126 SVTLQGEDNDGDCGVSYYVYKIFAGESETDRTHRSVTYTGIRKIQAPFTKGOQOQPC 125
 DB 126 SVTLQGEDNDGDCGVSYYVYKIFAGESETDRTHRSVTYTGIRKIQAPFTKGOQOQPC 125
 QY 186 VKRDMFLSPGELLEVTLKQVLYHGERIGVNICIRNNSNMVYKIKAMVQGVVYLFQ 245
 DB 186 TTRQFLMSDPKHLHLSLDEKITYHGEPISVNHNVTNNKTKYKIKISVQYADICLFN 245
 QY 246 NGSYRNTVASLETSECPYIOPGSSLOKMYLTPLLSNKRGRGIALDQIKRQOCLAST 245
 DB 246 TAQKCPVAMEAD--TVAPSSYFCVYTLTFLANNREKGLALDQIKRQOCLAST 245
 QY 306 TLLAOPDQDAFGVYISYAVKVLFL--GALGELISA-----ELPVLAMPKRGTK--A 355
 DB 304 TLLRSGANREILGITYSVKVLVVSRGQLGLDGLASSDAVELPFTLMHPKKEPPHR 363
 QY 356 KYIHADSOADVETFRQDTID 375
 DB 364 EVPEHETPVDTMLIBLDTND 383

RESULT 9
 ; US-10-038-010-52
 ; Sequence 52, Application US/10038010
 ; Patent No. US20030040089A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HYBRIGENICS
 ; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
 ; FILE REFERENCE: B4767A
 ; CURRENT APPLICATION NUMBER: US/10/038,010
 ; PRIOR FILING DATE: 2002-07-23
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Version 3.1
 ; SEQ ID NO 52
 ; LENGTH: 452
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: beta Arrestin 1
 ; LOCATION: (1)..(452)
 ; OTHER INFORMATION:
 ; US-10-038-010-52

Query Match
 Best Local Similarity 45.8%; Score 903.5; DB 14; Length 452;
 Matches 179; Conservative 76; Mismatches 112; Indels 13; Gaps 5;

QY 6 KYFKKCAPNGKVTLYMGKRDVDFVHSGVPEPIDGIVLDDERYIRDNKRVFGQIVCSFRYGR 65
 DB 7 RVFKKASPNKGLTVYGRDFVDHIDLVBPVGVVLDPEYLKE -RRVYVLTCAFRYGR 65
 QY 66 EDEVMGLNFOKELCLASEQIYPRPKSDKEQTKLOERLLKGSNAIPFTFNISPNAPS 125
 DB 66 EDLDVGLTFRKDLFVAVNQSFPAPEDKKPLTRLOERLLKGSNAIPFTFNISPNAPS 125
 QY 126 SVTLQGEDNDGDCGVSYYVYKIFAGESETDRTHRSVTYTGIRKIQAPFTKGOQOQPC 125
 DB 126 SVTLQGEDNDGDCGVSYYVYKIFAGESETDRTHRSVTYTGIRKIQAPFTKGOQOQPC 125

```
Db      126 SVTLQGPEDTGAACGVDEYKAFCAENLEEKHKRNSVRLVIRKVOYAPERPGOPTAE 185
Qy      186 VRKDFMLSPGELAEVLTLDKQLYLHGERIGVNICIRNNSNMKWKIKAMVOQGVVVLFQ 245
Db      186 TTRQFLMSDKRLHLEASIDKEIYHGERISVNVHTNTKTKVKKISVRODYDCLFN 245
Qy      246 NGSYRNTVASLETSEGCPIQPGSSLOQVMYLTPLSSNKKORGIALDQIKRQDCLAST 305
Db      246 TAQYKCPVAMEEADD--TVAPSSIFCKVYTLTPFLANNREKRGALDCKLKHEDTNLASS 303
Qy      306 TLLAQPDQDRAFGVITISYAVKVKLF--GALGELS-----ELPVLMPKPGTK--A 355
Db      304 TLLREGANREILGIIVSYKVKLVSRGGLGLDGLASDVAVELPFLMPKPEBPPHR 363
Qy      356 KVIHADSOADVETPROPTID 375
Db      364 EVPEHETPVDTNLIELDTND 383
```

```
RESULT 10
US-10-043-487-348
; Sequence 348, Application US/10043487
; Publication No. US2003005522041
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: PIERRE, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptides
; TITLE OF INVENTION: mammalian polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: Patent version 3.1
; SEQ ID NO 348
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Shigella flexneri
US-10-043-487-348
```

```
Query Match      45.6%; Score 900.5; DB 14; Length 418;
Best Local Similarity 46.8%; Pred. No. 1.7e-75;
Matches 178; Conservative 77; Mismatches 112; Indels 13; Gaps 5;

Qy      6 KVFKKCAPNGKVTLYMGKRDVPHVSGVEPIDGIIVLDDERYIRDNKRVFGQIVCSFRYGR 65
Db      7 RVFKKASPNKCLTVYLGKRDVPHIDLVPDGVVLVDPEYLKE--RRVYVTLTCAFRYGR 65
Qy      66 EEDVWGLNFQKELCLASEQIYPRPEKSDKEQTKLERLKKLGSNAIPTFNISPNAPS 125
Db      66 EDDDVGLTRKDLFVAVNVSFPAPEDKKELRLQERLKKLGEHAYPTTFELIPVLPFC 125
Qy      126 SVTLQGEDNDGPGCVSYVYKIFAGESETDRTHRSVTVLGIRKIQFARTKQGOQDCTL 185
Db      126 SVTLQGPEDTGAACGVDEYKAFCAENLEEKHKRNSVRLVIRKVOYAPERPGOPTAE 185
Qy      186 VRKDFMLSPGELAEVLTLDKQLYLHGERIGVNICIRNNSNMKWKIKAMVOQGVVVLFQ 245
Db      186 TTRQFLMSDKRLHLEASIDKEIYHGERISVNVHTNTKTKVKKISVRODYDCLFN 245
Qy      246 NGSYRNTVASLETSEGCPIQPGSSLOQVMYLTPLSSNKKORGIALDQIKRQDCLAST 305
Db      246 TAQYKCPVAMEEADD--TVAPSSIFCKVYTLTPFLANNREKRGALDCKLKHEDTNLASS 303
Qy      306 TLLAQPDQDRAFGVITISYAVKVKLF--GALGELS-----ELPVLMPKPGTK--A 355
Db      304 TLLREGANREILGIIVSYKVKLVSRGGLGLDGLASDVAVELPFLMPKPEBPPHR 363
Qy      356 KVIHADSOADVETPROPTID 375
Db      364 EVPEHETPVDTNLIELDTND 383
```

```
RESULT 11
US-09-880-137-2
; Sequence 2, Application US/09880137
; Patent No. US20020031295A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
; FILE REFERENCE: NMI-131
; CURRENT APPLICATION NUMBER: US/09/880,137
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-137-2
```

```
Query Match      45.5%; Score 897.5; DB 9; Length 418;
Best Local Similarity 46.8%; Pred. No. 3.3e-75;
Matches 178; Conservative 76; Mismatches 113; Indels 13; Gaps 5;

Qy      6 KVFKKCAPNGKVTLYMGKRDVPHVSGVEPIDGIIVLDDERYIRDNKRVFGQIVCSFRYGR 65
Db      7 RVFKKASPNKCLTVYLGKRDVPHIDLVPDGVVLVDPEYLKE--RRVYVTLTCAFRYGR 65
Qy      66 EEDVWGLNFQKELCLASEQIYPRPEKSDKEQTKLERLKKLGSNAIPTFNISPNAPS 125
Db      66 EDDDVGLTRKDLFVAVNVSFPAPEDKKELRLQERLKKLGEHAYPTTFELIPVLPFC 125
Qy      126 SVTLQGEDNDGPGCVSYVYKIFAGESETDRTHRSVTVLGIRKIQFARTKQGOQDCTL 185
Db      126 SVTLQGPEDTGAACGVDEYKAFCAENLEEKHKRNSVRLVIRKVOYAPERPGOPTAE 185
Qy      186 VRKDFMLSPGELAEVLTLDKQLYLHGERIGVNICIRNNSNMKWKIKAMVOQGVVVLFQ 245
Db      186 TTRQFLMSDKRLHLEASIDKEIYHGERISVNVHTNTKTKVKKISVRODYDCLFN 245
Qy      246 NGSYRNTVASLETSEGCPIQPGSSLOQVMYLTPLSSNKKORGIALDQIKRQDCLAST 305
Db      246 TAQYKCPVAMEEADD--TVAPSSIFCKVYTLTPFLANNREKRGALDCKLKHEDTNLASS 303
Qy      306 TLLAQPDQDRAFGVITISYAVKVKLF--GALGELS-----ELPVLMPKPGTK--A 355
Db      304 TLLREGANREILGIIVSYKVKLVSRGGLGLDGLASDVAVELPFLMPKPEBPPHR 363
Qy      356 KVIHADSOADVETPROPTID 375
Db      364 EVPEHETPVDTNLIELDTND 383

RESULT 12
US-09-800-137A-2
; Sequence 2, Application US/09800137A
; Publication No. US20030157553A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
; FILE REFERENCE: NMI-131
; CURRENT APPLICATION NUMBER: US/09/800,137A
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 418
; TYPE: PRT
```

Thu Feb 10 14:37:08 2005

ORGANISM: Homo sapiens
US-09-800-137A-2

Query Match

Best Local Similarity 45.5%; Score 897.5; DB 10; Length 418;
Matches 178; Conservative 76; Mismatches 113; Indels 13; Gaps 5;

Db 6 KVRKCAPNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDDEYIRDNKRYGQIVCSFRYGR 65
Qy 7 RVFKKASPNGLIYVLRGDFVHIDLVDPDGVVLDPEYLKE--RKYVYVLTLCARFGR 65
Db 6 EDEVMGLNFQKELCLASEQIYPRPEKSDKEQTKLOERLLKGSNAIPFTFNISNAPS 125
Qy 6 EDDVIGLITRKQDLFANVQSFPPAPDKKPLRLQERLLKQDEHAYPFTFEIPNLPFC 125
Db 126 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 185
Qy 126 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 185
Db 186 VKDPMLSPELELYTLDKQYLHGERIGVNICINNSKNKVKIKRAMYQGGVDVVLFPQ 245
Qy 186 TTRQPLMSDKPLHLASLDKELIYHGPISVNVHNTNNTVKKIKISTROYADICLFN 245
Db 246 NGSYRNTVASLSTSGCPIQSGSLQKVMYTLPLISSKQKRGIALDQGIKQDQCLAST 305
Qy 246 TAQYKCPVAOLEQD--TVAPSTFCVKYITLTPFLANNREKRGALDQGLHEDTNLASS 303
Db 304 TLLAOPDORDAFGVIISYVAVKYL--LGALGSELAS----ELPPVLMHKKQTK--A 355
Qy 304 TLLAOPDORDAFGVIISYVAVKYL--LGALGSELAS----ELPPVLMHKKQTK--A 355
Db 356 KVIHADSOADVETFFQDITD 375
Qy 364 EYENETPVDTNLTLELTDND 383

RESULT 13
US-09-880-137-8
Sequence 8, Application US/09880137
Patent No. US20020031295A1
GENERAL INFORMATION:
APPLICANT: Berstein, Gabriel
TITLE OF INVENTION: METHODS OF ASSAYING FOR G
FILE REFERENCE: NMI-131
CURRENT APPLICATION NUMBER: US/09/880,137
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FASTSEQ for Windows Version 4.0
LENGTH: 382
TYPE: PRY
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-880-137-8

Query Match
Best Local Similarity 44.9%; Score 885; DB 9; Length 382;
Matches 167; Conservative 77; Mismatches 98; Indels 4; Gaps 3;

Db 6 KVRKCAPNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDDEYIRDNKRYGQIVCSFRYGR 65
Qy 8 RVFKKASPNGLIYVLRGDFVHIDLVDPDGVVLDPEYLKE--RKYVYVLTLCARFGR 66
Db 6 EDEVMGLNFQKELCLASEQIYPRPEKSDKEQTKLOERLLKGSNAIPFTFNISNAPS 125
Qy 6 EDDVIGLITRKQDLFANVQSFPPAPDKKPLRLQERLLKQDEHAYPFTFEIPNLPFC 126
Db 126 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 185
Qy 126 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 185

us-10-056-405-2.rapb

Db

Qy 127 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 186

Db 186 VKDPMLSPELELYTLDKQYLHGERIGVNICINNSKNKVKIKRAMYQGGVDVVLFPQ 245

Qy 187 TTRHFLMSDKPLHLASLDKELIYHGPISVNVHNTNNTVKKIKISTROYADICLFN 245

Db 246 NGSYRNTVASLSTSGCPIQSGSLQKVMYTLPLISSKQKRGIALDQGIKQDQCLAST 305

Qy 247 TAQYKCPVAOLEQD--QVSPSTFCVKYITLTPFLANNREKRGALDQGLHEDTNLASS 304

Db 306 TLLAOPDORDAFGVIISYVAVKYL--LGALGSELASAEELPPVLMHKKP 351

Qy 305 TIVEGANKREVLGLIVSYVAVKLV--SRGDSVVELPVLHMKP 349

RESULT 14
US-09-800-137A-8
Sequence 8, Application US/09800137A
Patent No. US20030157553A1
GENERAL INFORMATION:
APPLICANT: Berstein, Gabriel
TITLE OF INVENTION: METHODS OF ASSAYING FOR G
FILE REFERENCE: NMI-131
CURRENT APPLICATION NUMBER: US/09/800,137A
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FASTSEQ for Windows Version 4.0
LENGTH: 382
TYPE: PRY
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-800-137A-8

Query Match
Best Local Similarity 44.9%; Score 885; DB 10; Length 382;
Matches 167; Conservative 77; Mismatches 98; Indels 4; Gaps 3;

Db 6 KVRKCAPNGKVTLYMGKRDVDFVHSGVEPIDGIVLDDDEYIRDNKRYGQIVCSFRYGR 65
Qy 8 RVFKKASPNGLIYVLRGDFVHIDLVDPDGVVLDPEYLKE--RKYVYVLTLCARFGR 66
Db 6 EDEVMGLNFQKELCLASEQIYPRPEKSDKEQTKLOERLLKGSNAIPFTFNISNAPS 125
Qy 6 EDDVIGLITRKQDLFANVQSFPPAPDKKPLRLQERLLKQDEHAYPFTFEIPNLPFC 126
Db 126 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 185
Qy 126 SVTLQGGEDNDGPGGVYVYKIFAGESETDRTHRSYVTLGIRKIQFAPYKGGQOCPCTL 185
Db 186 VKDPMLSPELELYTLDKQYLHGERIGVNICINNSKNKVKIKRAMYQGGVDVVLFPQ 245
Qy 187 TTRHFLMSDKPLHLASLDKELIYHGPISVNVHNTNNTVKKIKISTROYADICLFN 245
Db 246 NGSYRNTVASLSTSGCPIQSGSLQKVMYTLPLISSKQKRGIALDQGIKQDQCLAST 305
Qy 247 TAQYKCPVAOLEQD--QVSPSTFCVKYITLTPFLANNREKRGALDQGLHEDTNLASS 304
Db 306 TLLAOPDORDAFGVIISYVAVKYL--LGALGSELASAEELPPVLMHKKP 351
Qy 305 TIVEGANKREVLGLIVSYVAVKLV--SRGDSVVELPVLHMKP 349

RESULT 15
US-09-880-137-4
Sequence 4, Application US/09880137
Patent No. US20020031295A1
GENERAL INFORMATION:
APPLICANT: Berstein, Gabriel

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:45:38 / Search time 165 Seconds
(without alignments)
897.753 Million cell updates/sec

Title: US-10-056-405-2

Perfect score: 383

Sequence: 1 MYVNFKVFKKCAPNGKVTLY.....ADVETPRQDTIDQASVDPE 383

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searchod: 2105692 seqe, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	383	100.0	383	5	ABP52833 Anopheles
2	383	100.0	383	7	ABR84476 Mosquito
3	20	5.2	364	4	ABB61736 Drosophila
4	20	5.2	364	8	ADQ89588 Antagonist
5	11	2.9	382	4	AAQ67785 An exempl
6	11	2.9	398	7	ABR84484 Mosquito
7	11	2.9	401	4	ABB61964 Drosophila
8	11	2.9	409	4	AAQ67784 An exempl
9	11	2.9	409	4	AAQ67781 Amino aci
10	11	2.9	409	7	ADQ89583 Human pro
11	11	2.9	409	7	ADQ89581 Human pro
12	11	2.9	409	7	ADQ89581 Human pro
13	11	2.9	409	7	ADQ89581 Human pro
14	11	2.9	409	7	ADQ89581 Human pro
15	11	2.9	409	7	ADQ89581 Human pro
16	11	2.9	409	7	ADQ89581 Human pro
17	11	2.9	409	7	ADQ89581 Human pro
18	11	2.9	409	7	ADQ89581 Human pro
19	9	2.3	388	4	AAQ67782 Amino aci
20	9	2.3	388	4	AAQ67783 Amino aci
21	9	2.3	388	4	AAQ67783 Amino aci
22	8	2.1	22	7	ADQ89581 Human pro
23	8	2.1	22	7	ADQ89581 Human pro
24	8	2.1	22	7	ADQ89581 Human pro
25	8	2.1	22	7	ADQ89581 Human pro

26	8	2.1	38	4	ABB25470 Protein #
27	8	2.1	38	4	AAW75381 Human bon
28	8	2.1	38	4	AAW62570 Human bra
29	8	2.1	38	4	ABG57138 Human liv
30	8	2.1	38	5	ABG44975 Human pep
31	8	2.1	180	6	ADA00634 Human bet
32	8	2.1	182	7	ADD27419 Human adl
33	8	2.1	182	7	ADD27127 Human adl
34	8	2.1	182	7	ADD27197 Human adl
35	8	2.1	187	6	ABU70623 Human adl
36	8	2.1	187	7	ADD27507 Human adl
37	8	2.1	204	2	AAW20559 Helicobac
38	8	2.1	204	2	AAW24688 H. pylori
39	8	2.1	212	4	AAQ73899 Human col
40	8	2.1	279	7	ABO78666 Pseudomon
41	8	2.1	315	6	ADA00636 Human bet
42	8	2.1	360	6	ADA00635 Human bet
43	8	2.1	366	2	AAW20643 H. pylori
44	8	2.1	369	4	ABG13574 Novel hum
45	8	2.1	405	4	AAW78763 Human pro

ALIGNMENTS

RESULT 1
ABP52833 standard; protein: 383 AA.
ID ABP52833
XX
AC ABP52833:
XX
DT 01-NOV-2002 (first entry)
XX
DE Anopheles gambiae araeatin 1 protein seq ID NO:2.
XX
KW Anopheles gambiae; mosquito; olfactory gene; araeatin 1; pest control;
XX
KM odourant receptor; olfaction.
XX
OS Anopheles gambiae.
XX
PN W0200259274-A2.
XX
PD 01-AUG-2002.
XX
PF 26-JAN-2002; 2002MO-US002549.
XX
PR 26-JAN-2001; 2001US-0264649P.
XX
PR 24-JAN-2002; 2002US-00056405.
XX
PA (UYVA-) UNIV VANDERBILT.
XX
XX Zwiebel LJ;
XX WPI; 2002-627421/67.
XX N-PSDB; ABQ75102.
XX
PT New mosquito olfaction polypeptides and polynucleotides, useful for
PT mosquito management, i.e. controlling the pest and disease vectors, or
PT for identifying pest control agents.
XX
PS Claim 18; Fig 2; 96pp; English.
XX
XX The present invention describes a purified Anopheles gambiae olfaction
XX polypeptide comprising a 383, 394, 380, 411, 412, 391, 157 or 401 residue
XX amino acid sequence (see ABP52833 to ABP52840) (S1), a conservatively
XX modified amino acid sequence of them, or a sequence of (S1) with at least
XX 20 consecutive residues. Also described: (1) an isolated polynucleotide
XX comprising: (a) a nucleotide sequence encoding the purified Anopheles
XX gambiae olfaction polypeptide; or (b) a nucleotide sequence that
XX hybridises under stringent conditions to a hybridisation probe comprising
XX a 1964, 1239, 1142, 1236, 1194, 1176, 474 or 1206 nucleotide sequence
XX (see ABQ75102 to ABQ75105 and ABQ75110 to ABQ75113) (S2), or its
XX complement; and (2) a method for identifying an agent that binds to

CC mosquito olfaction molecules comprising: (a) providing an isolated
 CC mosquito olfaction molecule; (b) contacting a test agent with the
 CC isolated mosquito olfaction molecule; and (c) detecting specific binding
 CC of the test agent to the isolated mosquito olfaction molecule, where the
 CC presence of specific binding identifies the test agent as a mosquito
 CC olfaction-binding compound. The mosquito olfaction molecule are useful
 CC for mosquito management, i.e. controlling this pest and disease vector.
 CC method from the present invention of screening for substances that
 CC modulate arrestin-oligant receptor interaction is useful for identifying
 CC pest control agents. The present sequence represents Anopheles gambiae
 CC arrestin 1 from the present invention

Sequence 383 AA;

Query Match

Best Local Similarity 100.0%; Score 383; DB 5; Length 383;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYNFKVFKKCAPNGKVTLYMGRDVFVDHVGVEPIDGIYVLDEYIRDNKRVFGQIVCS 60
 DB 1 MYNFKVFKKCAPNGKVTLYMGRDVFVDHVGVEPIDGIYVLDEYIRDNKRVFGQIVCS 60
 QY 61 FRYGREDEVMGYNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKGNSNAIPPTNIS 120
 DB 61 FRYGREDEVMGYNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKGNSNAIPPTNIS 120
 QY 121 PNAPSSVTLQOGEEDNDGPGCVSYVYKIFAGESETRDRTHRSYTLGIRKIQAPPTKQO 180
 DB 121 PNAPSSVTLQOGEEDNDGPGCVSYVYKIFAGESETRDRTHRSYTLGIRKIQAPPTKQO 180
 QY 181 QPCTLVKRDQFMLSPELELEVTLDKQLYLHGERIGVNICIRNNSNNKVKIKAMVQOQVD 240
 DB 181 QPCTLVKRDQFMLSPELELEVTLDKQLYLHGERIGVNICIRNNSNNKVKIKAMVQOQVD 240
 QY 241 VVLFONGSYRNTVAASLETSEGCPIOPGSSLOKVMYLTPLSSNKKORGIALDQIKRQO 300
 DB 241 VVLFONGSYRNTVAASLETSEGCPIOPGSSLOKVMYLTPLSSNKKORGIALDQIKRQO 300
 QY 301 CLASTTLLAQDPORDAFGVIISYAVKVKLFGLAGELSAELPFVLMHPKRGTKAKVITHA 360
 DB 301 CLASTTLLAQDPORDAFGVIISYAVKVKLFGLAGELSAELPFVLMHPKRGTKAKVITHA 360
 QY 361 DSOADVETFRQDTIDQASVDPE 383
 DB 361 DSOADVETFRQDTIDQASVDPE 383

RESULT 2

ABR84476

ID ABR84476 standard; protein; 383 AA.
 AC ABR84476;
 XX

DT 15-JAN-2004 (first entry)

DE Mosquito olfaction molecule, arrestin 1.

KW Arrestin 1; mosquito; olfaction; insecticide; antimalarial.

OS Anopheles gambiae.

XX WO2003076590-A2.

PD 18-SEP-2003.

PF 10-MAR-2003; 2003WO-US007174.

FR 08-MAR-2002; 2002US-00094240.

PA (UYVA-) UNIV VANDERBILT.

XX Zwiebel LJ;

PI

XX

DR WPI; 2003-72231/68.
 DR N-FSDB; ACE79716.
 PT New mosquito arrestin 1 and 2 genes and polypeptides, useful for
 PT identifying mosquito olfaction molecule binding compounds which reduce
 XX the ability of mosquitoes to locate sources of bloodmeal, e.g. humans.
 PS Claim 1; Fig 2; 101pp; English.

CC The present sequence is the protein sequence of arrestin 1, a novel
 CC mosquito olfaction molecule. The invention provides 9 novel mosquito
 CC polypeptides and the nucleic acids encoding them. These are arrestin
 CC in a ligand-induced signal transduction pathway for the activation of
 CC mosquito olfaction. Arrestin functions to inhibit the activated signal
 CC and arrestin as an 'off' switch for the odorant receptors act as an 'on' switch,
 CC with the operation of the mosquito olfactory system, particularly the
 CC compounds that modulate arrestin 2 activity. These are useful for the
 CC control of mosquitoes, particularly by reducing their ability to locate
 CC sources of bloodmeal

Sequence 383 AA;

Query Match

Best Local Similarity 100.0%; Score 383; DB 7; Length 383;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYNFKVFKKCAPNGKVTLYMGRDVFVDHVGVEPIDGIYVLDEYIRDNKRVFGQIVCS 60
 DB 1 MYNFKVFKKCAPNGKVTLYMGRDVFVDHVGVEPIDGIYVLDEYIRDNKRVFGQIVCS 60
 QY 61 FRYGREDEVMGYNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKGNSNAIPPTNIS 120
 DB 61 FRYGREDEVMGYNFQKELCLASEQIYPRPEKSDKEQTKLQERLLKKGNSNAIPPTNIS 120
 QY 121 PNAPSSVTLQOGEEDNDGPGCVSYVYKIFAGESETRDRTHRSYTLGIRKIQAPPTKQO 180
 DB 121 PNAPSSVTLQOGEEDNDGPGCVSYVYKIFAGESETRDRTHRSYTLGIRKIQAPPTKQO 180
 QY 181 QPCTLVKRDQFMLSPELELEVTLDKQLYLHGERIGVNICIRNNSNNKVKIKAMVQOQVD 240
 DB 181 QPCTLVKRDQFMLSPELELEVTLDKQLYLHGERIGVNICIRNNSNNKVKIKAMVQOQVD 240
 QY 241 VVLFONGSYRNTVAASLETSEGCPIOPGSSLOKVMYLTPLSSNKKORGIALDQIKRQO 300
 DB 241 VVLFONGSYRNTVAASLETSEGCPIOPGSSLOKVMYLTPLSSNKKORGIALDQIKRQO 300
 QY 301 CLASTTLLAQDPORDAFGVIISYAVKVKLFGLAGELSAELPFVLMHPKRGTKAKVITHA 360
 DB 301 CLASTTLLAQDPORDAFGVIISYAVKVKLFGLAGELSAELPFVLMHPKRGTKAKVITHA 360
 QY 361 DSOADVETFRQDTIDQASVDPE 383
 DB 361 DSOADVETFRQDTIDQASVDPE 383

RESULT 3

ABB61736

ID ABB61736 standard; protein; 364 AA.
 AC ABB61736;
 XX

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 12000.

KW Drosophila; developmental biology; cell signalling; insecticide;

XX Drosophila melanogaster.

PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PP 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL05839.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 12000; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU6176-ABU30511), expressed DNA
 CC sequences (ABU01840-ABU16175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 364 AA;

Query Match 5.2%; Score 20; DB 4; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 VKKIKAMVQGVVLPONG 247
 |||||
 Db 227 VKKIKAMVQGVVLPONG 246

RESULT 4
 ID ADO89588 standard; protein; 364 AA.
 XX
 AC ADO89588;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Antagonist of cell cycle progression polypeptide #9.
 XX
 KM Cytostatic; cancer; cell division cycle; mitosis; meiosis;
 KM cell cycle progression.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO2004063162-A2.
 XX
 PD 29-JUL-2004.
 XX
 PP 31-DEC-2003; 2003WO-GB005635.
 XX
 PR 10-JAN-2003; 2003US-0439123P.
 PR 06-MAY-2003; 2003US-0468402P.
 XX
 PA (CYCL-) CYCLACEL LTD.
 XX
 PI Glover D, Bell G, Frenz L, Midgley C;
 XX
 DR WPI; 2004-544089/52.

DR N-PSDB; ADO89587.
 XX
 PT New cell cycle progression genes and proteins for modulating cell cycle
 PT progression in cells, for preventing, treating or diagnosing cell
 PT proliferative diseases (e.g. cancer) or for identifying modulators of
 PT mitosis or meiosis.
 XX
 PS Claim 2; SEQ ID NO 18; 461pp; English.
 XX
 CC The present invention relates to a polynucleotide for preventing,
 CC treating or diagnosing a disease in an individual. The composition of the
 CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for
 CC diagnosing, preventing or treating diseases (e.g. cell proliferative
 CC diseases such as cancer) in an individual. These may also be used for
 CC identifying substances capable of binding to or modulating the function
 CC of the polypeptide, capable of affecting the function of the
 CC corresponding gene, or capable of inhibiting the cell division cycle or
 CC cell cycle progression, preferably mitosis and/or meiosis. The present
 CC sequence represents an antagonist of cell cycle progression protein
 CC sequence.

SQ Sequence 364 AA;

Query Match 5.2%; Score 20; DB 8; Length 364;
 Best Local Similarity 100.0%; Pred. No. 1.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 VKKIKAMVQGVVLPONG 247
 |||||
 Db 227 VKKIKAMVQGVVLPONG 246

RESULT 5
 ID AAG67785 standard; protein; 382 AA.
 XX
 AC AAG67785;
 XX
 DT 10-DEC-2001 (first entry)
 XX
 DE An exemplary phosphorylation-independent arrestin mutant.
 XX
 KM Arrestin; phosphorylation-independent arrestin mutant;
 KM G protein-coupled receptor; GPCR; GPCR ligand; retinitis pigmentosa;
 KM stationary night blindness; colour blindness; nephrogenic DI;
 KM isolated glucocorticoid deficiency; hyperfunctioning thyroid adenoma;
 KM familial hypocalcaemic hypercalcaemia; hyperparathyroidism;
 KM neurological disorder.
 XX
 OS Unidentified.
 XX
 PN WO200167106-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US007304.
 XX
 PR 03-MAR-2000; 2000US-0186706P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Berstein G;
 XX
 DR WPI; 2001-602637/68.
 XX
 PT Identifying a G protein-coupled receptor ligand, useful for treating e.g.
 PT retinitis pigmentosa, color blindness or neurological disorders, uses
 PT phosphorylation-independent arrestin mutants particularly suited for in
 PT vitro screening assays.
 XX
 PS Disclosure; Page 47; 47pp; English.
 XX
 CC The present sequence represents an exemplary phosphorylation-independent

CC arrestin mutant. Such mutants are used in screening assays to identify
 CC ligands and/or modulators of G protein-coupled receptors (GPCRs). A
 CC method for identifying a GPCR ligand comprises contacting a composition
 CC comprising the GPCR and a constitutively active arrestin mutant with a
 CC test compound; and determining the ability of the test compound to
 CC modulate binding of arrestin mutant to the GPCR, where modulation of
 CC binding indicates that the test compound is a GPCR ligand. The method is
 CC used for identifying potential ligands and/or modulators of GPCRs,
 CC particularly modulators, for use as human therapeutics. Modulators of
 CC GPCR may be used for treating patients having e.g. retinitis pigmentosa,
 CC glucocorticoid deficiency, colour blindness, nephrogenic DI, isolated
 CC hypoadrenocorticism, hyperparathyroidism and neuroendocrine disorders.
 CC The methods may be used for screening pluralities of test
 CC compounds (e.g. a small molecule library of compounds) or a composition
 CC containing a plurality of GPCRs. The methods may be used in screening
 CC assays for identification of natural and surrogate agonists of orphan
 CC GPCRs, and for identification of GPCR antagonists and/or agonists
 SQ Sequence 382 AA;

Query Match 2.9%; Score 11; DB 4; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 341 ELPEVLMHPKP 351
 Db 339 ELPEVLMHPKP 349

RESULT 6
 ABR84484

ABR84484 standard; protein; 398 AA.

AC ABR84484;

DT 15-JAN-2004 (first entry)

DE Mosquito olfaction molecule, arrestin 2.

KW Arrestin 2; mosquito; olfaction; insecticide; antimalarial.

OS Anopheles gambiae.

Key Location/Qualifiers

FT Misc-difference 152 /note="Encoded by AG"

PN WO2003076590-A2.

PD 18-SEP-2003.

PF 10-MAR-2003; 2003WO-US007174.

PR 08-MAR-2002; 2002US-00094240.

PA (UYVA-) UNIV VANDERBILT.

PI Zwiebel LJ;

DR WPI; 2003-722331/68.

DR N-PSDB; ACF79731.

PT New mosquito arrestin 1 and 2 genes and polypeptides, useful for
 PT identifying mosquito olfaction molecule binding compounds which reduce
 PT the ability of mosquitoes to locate sources of bloodmeal, e.g. humans.
 PS Claim 16; Fig 12b; 101pp; English.

CC The present sequence is the protein sequence of arrestin 2, a novel
 CC mosquito olfaction molecule. The invention provides 9 novel mosquito
 CC polypeptides and the nucleic acids encoding them. These are arrestin 1
 CC and 2 and odorant receptor molecules 1-7. The odorant receptors function

CC in a ligand-induced signal transduction pathway for the activation of
 CC mosquito olfaction. Arrestin functions to inhibit the activated signal
 CC transduction cascade. Thus, the odorant receptors act as an 'on' switch,
 CC and arrestin as an 'off' switch for the odorant detection system of the
 CC mosquito. Methods are provided for identifying compounds that interfere
 CC with the operation of the mosquito olfactory system, particularly
 CC compounds that modulate arrestin 2 activity. These are useful for the
 CC control of mosquitoes, particularly by reducing their ability to locate
 CC sources of bloodmeal
 SQ Sequence 398 AA;

Query Match 2.9%; Score 11; DB 7; Length 398;
 Best Local Similarity 100.0%; Pred. No. 0.1;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 RYGREDEVMG 72
 Db 61 RYGREDEVMG 71

RESULT 7

ABR61964

ABR61964 standard; protein; 401 AA.

AC ABR61964;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 12684.

KW Drosophila; developmental biology; cell signalling; insecticide;

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

DR Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL06067.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 12684; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 401 AA;

Query Match 2.9%; Score 11; DB 4; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0.1;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 62 RYGREDEVMG 72
| | | | | | | | | |
DB 61 RYGREDEVMG 71

RESULT 8
AAG67784
ID AAG67784 standard; protein; 409 AA.

XX AAG67784;

DT 10-DEC-2001 (first entry)

DE An exemplary phosphorylation-independent arrestin mutant.

XX Arrestin; phosphorylation-independent arrestin mutant;

KW G protein-coupled receptor; GPCR; GPCR ligand; retinitis pigmentosa;
stationary night blindness; colour blindness; nephrogenic DI;

KW isolated glucocorticoid deficiency; hyperfunctioning thyroid adenoma;
familial hypocalcaemic hypercalcaemia; hyperparathyroidism;
neurological disorder.

XX Unidentified.

OS WO200167106-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US007304.

XX 03-MAR-2000; 2000US-0186706P.

XX (MILL-) MILLENNIUM PHARM INC.

PA Berstein G;

PI WPI; 2001-602637/68.

PT Identifying a G protein-coupled receptor ligand, useful for treating e.g.
retinitis pigmentosa, color blindness or neurological disorders, uses
phosphorylation-independent arrestin mutants particularly suited for in
vitro screening assays.

XX Disclosure; Page 47; 47pp; English.

XX The present sequence represents an exemplary phosphorylation-independent
arrestin mutant. Such mutants are used in screening assays to identify
ligands and/or modulators of G protein-coupled receptors (GPCRs). A
method for identifying a GPCR ligand comprises contacting a composition
comprising the GPCR and a constitutively active arrestin mutant with a
test compound; and determining the ability of the test compound to
modulate binding of arrestin mutant to the GPCR, where modulation of
binding indicates that the test compound is a GPCR ligand. The method is
used for identifying potential ligands and/or modulators of GPCRs,
particularly modulators, for use as human therapeutics. Modulators of
GPCR may be used for treating patients having e.g. retinitis pigmentosa,
stationary night blindness, colour blindness, nephrogenic DI, isolated
glucocorticoid deficiency, hyperfunctioning thyroid adenoma, familial
hypocalcaemic hypercalcaemia, hyperparathyroidism and neurological
disorders. The methods may be used for screening pluralities of test
compounds (e.g. a small molecule library for screening pluralities of test
compounds containing a plurality of GPCRs. The methods may be used in screening
assays for identification of natural and surrogate agonists of orphan
GPCRs, and for identification of GPCR antagonists and/or agonists

XX Sequence 409 AA;

Query Match 2.9%; Score 11; DB 4; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELPEVLNMPKP 351
| | | | | | | | | |
DB 339 ELPEVLNMPKP 349

RESULT 9
AAG67781
ID AAG67781 standard; protein; 409 AA.

XX AAG67781;

DT 10-DEC-2001 (first entry)

DE Amino acid sequence of human beta-arrestin 2.

XX Arrestin; phosphorylation-independent arrestin mutant;

KW G protein-coupled receptor; GPCR; GPCR ligand; retinitis pigmentosa;
stationary night blindness; colour blindness; nephrogenic DI;

KW isolated glucocorticoid deficiency; hyperfunctioning thyroid adenoma;
familial hypocalcaemic hypercalcaemia; hyperparathyroidism;
neurological disorder.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Region 156..185
FT /note="putative phosphorylation-recognition region"

XX WO200167106-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US007304.

XX 03-MAR-2000; 2000US-0186706P.

XX (MILL-) MILLENNIUM PHARM INC.

PA Berstein G;

PI WPI; 2001-602637/68.

PT Identifying a G protein-coupled receptor ligand, useful for treating e.g.
retinitis pigmentosa, color blindness or neurological disorders, uses
phosphorylation-independent arrestin mutants particularly suited for in
vitro screening assays.

XX Disclosure; Page 46; 47pp; English.

XX The present sequence represents an arrestin protein. The specification
describes phosphorylation-independent arrestin mutants. These mutants are
used in screening assays to identify ligands and/or modulators of G
protein-coupled receptors (GPCRs). A method for identifying a GPCR ligand
comprises contacting a composition comprising the GPCR and a
constitutively active arrestin mutant with a test compound; and
determining the ability of the test compound to modulate binding of
arrestin mutant to the GPCR, where modulation of binding indicates that
the test compound is a GPCR ligand. The method is used for identifying
potential ligands and/or modulators of GPCRs, particularly modulators,
for use as human therapeutics. Modulators of GPCR may be used for
treating patients having e.g. retinitis pigmentosa, stationary night
blindness, colour blindness, nephrogenic DI, isolated glucocorticoid
deficiency, hyperfunctioning thyroid adenomas, familial hypocalcaemic
hypercalcaemia, hyperparathyroidism and neurological disorders. The
methods may be used for screening pluralities of test compounds (e.g. a
small molecule library for screening pluralities of test compounds containing a
plurality of GPCRs. The methods may be used in screening assays for
identification of natural and surrogate agonists of orphan GPCRs, and for
identification of GPCR antagonists and/or agonists

XX Sequence 409 AA;

Query Match 2.9%; Score 11; DB 4; Length 409;

Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 ELFPVLMHPKP 351
Db 339 ELFPVLMHPKP 349

RESULT 10
ADES8183
ID ADES8183 standard; protein; 409 AA.
XX AC ADES8183;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein P32121, SEQ ID NO 4054.
XX DE Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GHEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M,
XX WPI: 2003-268312/26.
XX DR GENBANK; P32121.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PS preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition,
a method for identifying a method for identifying a compound which
specifies a method for identifying a compound which regulates the
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
therapy). The sequence presented is a human protein (shown in Table 2 of
the specification) which is differentially expressed during pain. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 409 AA;
XX

Query Match
Best Local Similarity 2.9%; Score 11; DB 7; Length 409;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 ELFPVLMHPKP 351
Db 339 ELFPVLMHPKP 349

RESULT 11
ADES8179
ID ADES8179 standard; protein; 409 AA.
XX AC ADES8179;
XX DT 29-JAN-2004 (first entry)
XX DE Human Protein P32121, SEQ ID NO 4050.
XX DE Human; pain; neuronal tissue; gene therapy;
XX KW spinal segmental nerve injury; chronic constriction injury; CCI;
XX KW spared nerve injury; SNI; Chung.
XX OS Homo sapiens.
XX PN WO2003016475-A2.
XX PD 27-FEB-2003.
XX PF 14-AUG-2002; 2002WO-US025765.
XX PR 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX PA (GHEO) GEN HOSPITAL CORP.
XX PA (FARB) BAYER AG.
XX PI Woolf C, D'urso D, Befort K, Costigan M,
XX WPI: 2003-268312/26.
XX DR GENBANK; P32121.
XX PT New composition comprising two or more isolated polypeptides, useful for
XX PS preparing a medicament for treating pain in an animal.
XX PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
or human polynucleotides or a polynucleotide which represents a fragment,
derivative or allelic variation of the nucleic acid sequence. Also
claimed are a vector comprising the novel polynucleotide, a host cell
comprising the vector, a method for identifying a nucleotide sequence
which is differentially regulated in an animal subjected to pain and a
kit to perform the method, an array, a method for identifying an agent
that increases or decreases the expression of the polynucleotide sequence
that is differentially expressed in neuronal tissue of a first animal
subjected to pain, a method for identifying a compound which regulates
the expression of a polynucleotide sequence which is differentially
expressed in an animal subjected to pain, a method for identifying a
compound that regulates the activity of one or more of the
polynucleotides, a method for producing a pharmaceutical composition,
a method for identifying a method for identifying a compound which
specifies a method for identifying a compound which regulates the
pain and a pharmaceutical composition comprising the one or more
polypeptides or their antibodies. The polynucleotide or the compound that
modulates its activity is useful for preparing a medicament for treating
pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 409 AA;

SO Query Match 2.9%; Score 11; DB 7; Length 409;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 341 ELPFVLMHPRK 351
 |||||
 Db 339 ELPFVLMHPRK 349

RESULT 12

ID ADN95251 standard; protein; 409 AA.

AC ADN95251;

DT 01-JUL-2004 (first entry)

DE Human BEC/LBC-related protein sequence SegID173.

XX growth; differentiation; blood endothelial cell; BEC;

KM lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;

KM lymphatic growth agent; VEGF-C; VEGF-D; angiogenic; cytoskeletal;

KM vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;

KM inflammatory disease; cancer metastasis; lymphatic system; human.

OS Homo sapiens.

PN WO2003080640-A1.

PD 02-OCT-2003.

PF 07-MAR-2003; 2003WO-US006900.

PR 07-MAR-2002; 2002US-0363019P.

PA (LUDM-) LUDWIG INST CANCER RES.

PA (LICN) LICENTIA LTD.

PI Alltalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;

DR WPI; 2003-876899/81.

DR N-PSDB; ADN95252.

PS Example 1; SEQ ID NO 173; 176pp; English.

CC The invention relates to a method of differentially modulating the
 CC growth or differentiation of blood endothelial cells (BEC) or lymphatic
 CC endothelial cells (LEC) comprises contacting endothelial cells with a
 CC composition comprising an agent that differentially modulates blood or
 CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises
 CC identifying a human subject with lymphoedema and with a mutation in at
 CC least one allele of a gene encoding a LEC protein, where the mutation
 CC correlates with lymphoedema in human subjects, and with the proviso that
 CC the LEC protein is not VEGFR-3; and administering to the subject a
 CC composition comprising a lymphatic growth agent selected from VEGF-C or
 CC VEGF-D polypeptides and polynucleotides. The invention may be useful for
 CC the development of compounds with an angiogenic, cytoskeletal,
 CC vasotropic or antiinflammatory activity or for gene therapy. The method
 CC is useful in modulating the growth or differentiation of blood
 CC endothelial cells or lymphatic endothelial cells, in treating hereditary
 CC lymphoedema, in screening for an endothelial cell disorder or
 CC predisposition to the disorder or in monitoring the efficacy or toxicity
 CC of a drug on endothelial cells. The agent is useful in manufacturing a
 CC medicament for the differential modulation of blood vessel endothelial

CC cell or lymphatic vessel endothelial cell growth or differentiation. The
 CC lymphatic growth agent may also be used in manufacturing a medicament for
 CC the treatment of hereditary lymphoedema resulting from a mutation in a
 CC LEC gene or of other diseases involving the lymphatic vessels, such as
 CC various inflammatory diseases and cancer metastasis via the lymphatic
 CC system. The present sequence is that of a human BEC/LBC differentially
 CC expressed protein which is related to the method of the invention. Note:
 CC This sequence does not appear in the specification but was obtained by
 CC the indexer using the source data given in table 14 of the specification.

XX Sequence 409 AA;

SO Query Match 2.9%; Score 11; DB 7; Length 409;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 341 ELPFVLMHPRK 351
 |||||
 Db 339 ELPFVLMHPRK 349

RESULT 13

ID ABG69496 standard; protein; 410 AA.

AC ABG69496;

DT 21-OCT-2002 (first entry)

DE Rat bait protein beta arrestin 2.

XX Rat; Yeast two-hybrid assay; adipocyte; bait protein; NIDDM;

KM non-insulin diabetes mellitus; obesity; selected interacting domain; SID;

KM protein-protein interaction map; PTM; anorectic; metabolic disorder.

OS Rattus sp.

PN WO200253726-A2.

PD 11-JUL-2002.

PF 28-DEC-2001; 2001WO-EP015423.

PR 02-JAN-2001; 2001US-0259377P.

PA (HYBR-) HYBRIGENICS.

PA (CNRS) CENT NAT RECH SCI.

PI Legrain P, Marullo S, Jockers R;

DR WPI; 2002-583612/62.

DR N-PSDB; ABG51032.

PS Novel complex of protein-protein interactions in adipocyte cells for
 PT identifying compounds that modulate the protein-protein interactions and
 PT useful for treating obesity and metabolic disorders.

PS Claim 1; Page 53; 125pp; English.

CC The invention relates to a complex of protein-protein interactions
 CC (forming a protein-protein interaction map, PM) in adipocyte cells as
 CC defined in the specification, or polynucleotides in adipocytes encoding
 CC for the polypeptides. Also included are a recombinant cell expressing the
 CC interacting polypeptides and a method of selecting a modulating compound
 CC in adipocyte cells, by cultivating a recombinant host cell on a selective
 CC medium containing a modulating compound and a reporter gene the
 CC expression of which is toxic for the recombinant host cell which is
 CC transformed with two vectors, where the first vector comprises a
 CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain
 CC and the second vector comprising a polynucleotide encoding a second
 CC hybrid polypeptide and an activating domain that activates the toxic
 CC reporter gene, when the first and second hybrid polypeptides interact and
 CC selecting the modulating compound which inhibits the growth of the

CC recombinant host cell (i.e. using the yeast two-hybrid system). The
CC complexes are useful for identifying compounds that modulate the protein-
CC protein interactions and useful for treating obesity and metabolic
CC disorders e.g. non-insulin dependent diabetes mellitus, and metabolic
CC obesity or metabolic diseases. The interactions between the protein of
CC the complex further define a set of selected interacting domains, STD.
CC The present sequence represents a member of the protein complex of the
CC invention, used as the bait protein in the yeast two- hybrid assay
SQ Sequence 410 AA;

Query Match
Best Local Similarity 2.9%; Score 11; DB 5; Length 410;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 ELFPVLMHKKP 351
DB 340 ELFPVLMHKKP 350

RESULT 14

AD58177
ID ADE58177 standard; protein; 410 AA.
AC ADE58177;
XX
XX

DT 29-JAN-2004 (first entry)
DE Rat Protein P29067, SEQ ID NO 4048.
XX
XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX

OS Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX

PF 14-AUG-2002; 2002WO-US025765.
XX
XX

PR 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX

PA (GHEO) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.
XX

PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; P29067.
XX

PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC claimed are a vector comprising the nucleic acid sequence. Also
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC polynucleotide, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC specification, but was obtained in electronic form directly from WIPO at
SQ Sequence 410 AA;

Query Match
Best Local Similarity 2.9%; Score 11; DB 7; Length 410;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 341 ELFPVLMHKKP 351
DB 340 ELFPVLMHKKP 350

RESULT 15

AD58181
ID ADE58181 standard; protein; 410 AA.
AC ADE58181;
XX
XX

DT 29-JAN-2004 (first entry)
DE Rat Protein P29067, SEQ ID NO 4052.
XX
XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX

OS Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX

PF 14-AUG-2002; 2002WO-US025765.
XX
XX

PR 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX

PA (GHEO) GEN HOSPITAL CORP.
XX
XX (FARB) BAYER AG.
XX

PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX
XX GENBANK; P29067.
XX

PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC claimed are a vector comprising the nucleic acid sequence. Also
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 410 AA;

Query Match 2.9%; Score 11; DB 7; Length 410;

Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 341 ELPPVLMHPPK 351
 |||||
 Db 340 ELPPVLMHPPK 350

Search completed: February 9, 2005, 23:58:30
 Job time : 166 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:55:44 / Search time 43 Seconds
(without alignments)
664.897 Million cell updates/sec

Title: US-10-056-405-2
Perfect score: 383
Sequence: 1 MVNFKVFKKCAPNGKVTLY.....ADVETFRQDTIDQASVDFE 383

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pdp: *
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pdp: *
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pdp: *
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pdp: *
5: /cgn2_6/ptodata/1/1aa/PTCUG_COMB.pdp: *
6: /cgn2_6/ptodata/1/1aa/backfill1.pdp: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	2.9	382	4 US-09-880-137-8	Sequence 8, Appli
2	11	2.9	409	4 US-09-880-137-4	Sequence 4, Appli
3	11	2.9	409	4 US-09-880-137-7	Sequence 7, Appli
4	9	2.3	388	4 US-09-880-137-5	Sequence 5, Appli
5	9	2.3	388	4 US-09-880-137-6	Sequence 6, Appli
6	8	2.1	279	4 US-09-252-991A-27414	Sequence 27414, A
7	8	2.1	315	4 US-09-949-016-9940	Sequence 9940, Ap
8	8	2.1	401	4 US-09-949-016-8383	Sequence 8383, Ap
9	8	2.1	401	4 US-09-949-016-8384	Sequence 8384, Ap
10	8	2.1	410	4 US-09-880-137-3	Sequence 3, Appli
11	8	2.1	418	4 US-09-880-137-1	Sequence 1, Appli
12	8	2.1	418	4 US-09-880-137-2	Sequence 2, Appli
13	8	2.1	418	4 US-08-484-905-28	Sequence 28, Appli
14	7	1.8	15	3 US-08-481-985B-28	Sequence 28, Appli
15	7	1.8	15	3 US-08-370-476-28	Sequence 28, Appli
16	7	1.8	15	3 US-08-992-877-8	Sequence 8, Appli
17	7	1.8	15	3 US-08-480-190-54	Sequence 54, Appli
18	7	1.8	16	2 US-08-488-379-54	Sequence 54, Appli
19	7	1.8	16	4 US-08-475-399A-54	Sequence 54, Appli
20	7	1.8	16	4 US-08-077-255A-54	Sequence 54, Appli
21	7	1.8	16	5 PCT-US93-07545-54	Sequence 54, Appli
22	7	1.8	19	4 US-09-423-097-4	Sequence 4, Appli
23	7	1.8	20	2 US-08-934-915-43	Sequence 43, Appli
24	7	1.8	20	2 US-08-934-915-81	Sequence 81, Appli
25	7	1.8	20	3 US-09-367-953B-28	Sequence 28, Appli
26	7	1.8	70	3 US-09-367-953B-29	Sequence 29, Appli
27	7	1.8	70	3 US-09-367-953B-30	Sequence 30, Appli

28	7	1.8	70	3	US-09-367-953B-31	Sequence 31, Appli
29	7	1.8	70	3	US-09-367-953B-52	Sequence 52, Appli
30	7	1.8	70	3	US-09-367-953B-55	Sequence 55, Appli
31	7	1.8	70	3	US-09-367-953B-65	Sequence 65, Appli
32	7	1.8	70	3	US-09-367-953B-68	Sequence 68, Appli
33	7	1.8	70	3	US-09-367-953B-70	Sequence 70, Appli
34	7	1.8	70	3	US-09-367-953B-71	Sequence 71, Appli
35	7	1.8	70	3	US-09-367-953B-72	Sequence 72, Appli
36	7	1.8	70	3	US-09-367-953B-74	Sequence 74, Appli
37	7	1.8	70	3	US-09-367-953B-75	Sequence 75, Appli
38	7	1.8	70	3	US-09-367-953B-76	Sequence 76, Appli
39	7	1.8	70	3	US-09-367-953B-78	Sequence 78, Appli
40	7	1.8	70	3	US-09-367-953B-80	Sequence 80, Appli
41	7	1.8	70	3	US-09-367-953B-82	Sequence 82, Appli
42	7	1.8	70	3	US-09-367-953B-83	Sequence 83, Appli
43	7	1.8	70	3	US-09-367-953B-84	Sequence 84, Appli
44	7	1.8	70	3	US-09-367-953B-85	Sequence 85, Appli
45	7	1.8	70	3	US-09-367-953B-86	Sequence 86, Appli

ALIGNMENTS

```
RESULT 1
US-09-880-137-8
; Sequence 8, Application US/09880137
; Patent No. 6640025
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: NMI-131
; CURRENT APPLICATION NUMBER: US/09/880,137
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-880-137-8

Query Match      2.9%; Score 11; DB 4; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      341 ELRPVLMHPRP 351
DB      339 ELRPVLMHPRP 349

RESULT 2
US-09-880-137-4
; Sequence 4, Application US/09880137
; Patent No. 6640025
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: NMI-131
; CURRENT APPLICATION NUMBER: US/09/880,137
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 409
; TYPE: PRT
```

/ ORGANISM: Homo sapiens

US-09-880-137-4

Query Match

Best Local Similarity 2.9%; Score 11; DB 4; Length 409;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Qy 341 ELFPVLMHMKP 351
Db 339 ELFPVLMHMKP 349

RESULT 3

US-09-880-137-7

/ Sequence 7, Application US/09880137

/ Patent No. 6640025

/ GENERAL INFORMATION:

/ APPLICANT: Berstein, Gabriel

/ TITLE OF INVENTION: METHODS OF ASSAYING FOR G

/ FILE REFERENCE: MNI-131

/ CURRENT APPLICATION NUMBER: US/09/880,137

/ PRIOR FILING DATE: 2001-03-05

/ PRIOR APPLICATION NUMBER: US 60/186,706

/ NUMBER OF SEQ ID NOS: 8

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 7

/ LENGTH: 409

/ TYPE: PRT

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation

US-09-880-137-7

Query Match

Best Local Similarity 2.9%; Score 11; DB 4; Length 409;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Qy 341 ELFPVLMHMKP 351
Db 339 ELFPVLMHMKP 349

RESULT 4

US-09-880-137-5

/ Sequence 5, Application US/09880137

/ Patent No. 6640025

/ GENERAL INFORMATION:

/ APPLICANT: Berstein, Gabriel

/ TITLE OF INVENTION: METHODS OF ASSAYING FOR G

/ FILE REFERENCE: MNI-131

/ CURRENT APPLICATION NUMBER: US/09/880,137

/ PRIOR FILING DATE: 2001-03-05

/ PRIOR APPLICATION NUMBER: US 60/186,706

/ NUMBER OF SEQ ID NOS: 8

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 5

/ LENGTH: 388

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-09-880-137-5

Query Match

Best Local Similarity 2.3%; Score 9; DB 4; Length 388;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Qy 22 GKRDVFDHV 30
Db 19 GKRDVFDHV 27

RESULT 5

US-09-880-137-6

/ Sequence 6, Application US/09880137

/ Patent No. 6640025

/ GENERAL INFORMATION:

/ APPLICANT: Berstein, Gabriel

/ TITLE OF INVENTION: METHODS OF ASSAYING FOR G

/ FILE REFERENCE: MNI-131

/ CURRENT APPLICATION NUMBER: US/09/880,137

/ PRIOR FILING DATE: 2001-03-05

/ PRIOR APPLICATION NUMBER: US 60/186,706

/ NUMBER OF SEQ ID NOS: 8

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 6

/ LENGTH: 388

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-09-880-137-6

Query Match

Best Local Similarity 2.3%; Score 9; DB 4; Length 388;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Qy 22 GKRDVFDHV 30
Db 19 GKRDVFDHV 27

RESULT 6

US-09-252-991A-27414

/ Sequence 27414, Application US/09252991A

/ Patent No. 6551795

/ GENERAL INFORMATION:

/ APPLICANT: Marc J. Rubenfield et al.

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

/ FILE REFERENCE: 107196.136

/ CURRENT APPLICATION NUMBER: US/09/252,991A

/ PRIOR FILING DATE: 1999-02-18

/ PRIOR APPLICATION NUMBER: US 60/074,788

/ PRIOR FILING DATE: 1998-02-18

/ PRIOR APPLICATION NUMBER: US 60/094,190

/ NUMBER OF SEQ ID NOS: 33142

/ SEQ ID NO 27414

/ LENGTH: 279

/ TYPE: PRT

/ ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27414

Query Match

Best Local Similarity 2.1%; Score 8; DB 4; Length 279;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;Qy 336 GELSAELP 343
Db 30 GELSAELP 37

RESULT 7

US-09-949-016-9940

/ Sequence 9940, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ FILE REFERENCE: C1001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ PRIOR FILING DATE: 2000-04-14

;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9940
;; LENGTH: 315
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-9940

Query Match 2.1%; Score 8; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GKRDPVDH 29
Db 24 GKRDPVDH 31

RESULT 8
US-09-949-016-8383
;; Sequence 8383, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; FILE REFERENCE: C0001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8383
;; LENGTH: 401
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-8383

Query Match 2.1%; Score 8; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GKRDPVDH 29
Db 6 GKRDPVDH 13

RESULT 9
US-09-949-016-8384
;; Sequence 8384, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; FILE REFERENCE: C0001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 8384
;; LENGTH: 401
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-8384

Query Match 2.1%; Score 8; DB 4; Length 401;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GKRDPVDH 29
Db 6 GKRDPVDH 13

RESULT 10
US-09-880-137-3
;; Sequence 3, Application US/09880137
;; Patent No. 6640025
;; GENERAL INFORMATION:
;; APPLICANT: Berstein, Gabriel
;; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
;; FILE REFERENCE: W01-131
;; CURRENT APPLICATION NUMBER: US/09/880,137
;; PRIOR FILING DATE: 2001-03-05
;; PRIOR APPLICATION NUMBER: US 60/186,706
;; PRIOR FILING DATE: 2000-03-03
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 410
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-880-137-3

Query Match 2.1%; Score 8; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GKRDPVDH 29
Db 23 GKRDPVDH 30

RESULT 11
US-09-880-137-1
;; Sequence 1, Application US/09880137
;; Patent No. 6640025
;; GENERAL INFORMATION:
;; APPLICANT: Berstein, Gabriel
;; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
;; FILE REFERENCE: W01-131
;; CURRENT APPLICATION NUMBER: US/09/880,137
;; PRIOR FILING DATE: 2001-03-05
;; PRIOR APPLICATION NUMBER: US 60/186,706
;; PRIOR FILING DATE: 2000-03-03
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 418
;; TYPE: PRT
;; ORGANISM: Bos taurus
US-09-880-137-1

Query Match 2.1%; Score 8; DB 4; Length 418;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 GKRDVVDH 29
Db 23 GKRDVVDH 30

RESULT 12
US-09-880-137-2
Sequence 2, Application US/09880137
Patent No. 6640025
GENERAL INFORMATION:
APPLICANT: Berstein, Gabriel
TITLE OF INVENTION: METHODS OF ASSAYING FOR G
FILE REFERENCE: NMI-131
CURRENT APPLICATION NUMBER: US/09/880,137
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 60/186,706
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 418
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-137-2

Query Match
Best Local Similarity 2.1%; Score 8; DB 4; Length 418;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22 GKRDVVDH 29
Db 23 GKRDVVDH 30

RESULT 13
US-08-484-905-28
Sequence 28, Application US/08484905
Patent No. 5976551
GENERAL INFORMATION:
APPLICANT: Motiez, Estelle
APPLICANT: Aabastado, Jean-Pierre
TITLE OF INVENTION: An Altered Major Histocompatibility
TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,905
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.

REGISTRATION NUMBER: 33,332
REFERENCE/DOCKET NUMBER: 03495.0106-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-484-905-28

Query Match
Best Local Similarity 1.8%; Score 7; DB 2; Length 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 256 LETSEGC 262
Db 1 LETSEGC 7

RESULT 14
US-08-481-985B-28
Sequence 28, Application US/08481985B
Patent No. 601144
GENERAL INFORMATION:
APPLICANT: Motiez, Estelle
APPLICANT: Aabastado, Jean-Pierre
TITLE OF INVENTION: Altered Major Histocompatibility Complex
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,985B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0106-04000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-481-985B-28

Query Match
1.8%; Score 7; DB 3; Length 15;

Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: February 10, 2005, 00:11:11
Job time : 44 secs

QY 256 LETSEGC 262
|||||
Db 1 LETSEGC 7

RESULT 15
US-08-370-476-28
Sequence 28, Application US/08370476
Patent No. 6153408
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abatado, Jean-Pierre
APPLICANT: Kourilsky, Philippe
APPLICANT: Lorie, Yu-Chun
APPLICANT: Ojcius, David
APPLICANT: Casrouge, Armada
TITLE OF INVENTION: Altered Major Histocompatibility Complex
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flannegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyere, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243.0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-370-476-28

Query Match 1.8%; Score 7; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 256 LETSEGC 262
|||||
Db 1 LETSEGC 7

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:58:40 ; Search time 131 Seconds
(without alignment)
952.245 Million cell updates/sec

Title: US-10-056-405-2
Perfect score: 383
Sequence: 1 MVYVFKYFKKCAPNGKVTLY.....ADVFTRPDRTDQASVDPE 383

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1373511 seqs, 325702437 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBSCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_PUBSCOMB.pep:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
 - 5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBSCOMB.pep:*
 - 6: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
 - 7: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
 - 8: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_PUBSCOMB.pep:*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBSCOMB.pep:*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBSCOMB.pep:*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBSCOMB.pep:*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_PUBSCOMB.pep:*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_PUBSCOMB.pep:*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBSCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	383	100.0	383	14 US-10-094-240-2	Sequence 2, Appl1
2	383	100.0	383	14 US-10-056-405-2	Sequence 2, Appl1
3	382	9	382	9 US-09-880-137-8	Sequence 8, Appl1
4	382	10	382	10 US-09-800-137A-8	Sequence 8, Appl1
5	398	14	398	14 US-10-094-240-25	Sequence 25, Appl1
6	401	14	401	14 US-10-094-240-27	Sequence 27, Appl1
7	409	9	409	9 US-09-880-137-4	Sequence 4, Appl1
8	409	9	409	9 US-09-880-137-7	Sequence 7, Appl1
9	409	10	409	10 US-09-800-137A-4	Sequence 4, Appl1
10	409	10	409	10 US-09-800-137A-7	Sequence 4, Appl1
11	410	14	410	14 US-10-038-010-54	Sequence 54, Appl1
12	466	16	466	16 US-10-722-357-31	Sequence 31, Appl1
13	2.3	388	9	US-09-880-137-5	Sequence 5, Appl1

14	9	2.3	388	9	US-09-880-137-6	Sequence 6, Appl1
15	9	2.3	388	10	US-09-800-137A-5	Sequence 5, Appl1
16	9	2.3	388	10	US-09-800-137A-6	Sequence 6, Appl1
17	8	2.1	48	9	US-09-864-761-40768	Sequence 40768, A
18	8	2.1	48	15	US-10-424-599-154582	Sequence 154582, A
19	8	2.1	153	15	US-10-424-599-148204	Sequence 148204, A
20	8	2.1	204	16	US-10-335-977-7580	Sequence 7580, Ap
21	8	2.1	208	16	US-10-767-701-41403	Sequence 41403, A
22	8	2.1	212	14	US-10-106-698-4673	Sequence 4673, Ap
23	8	2.1	225	15	US-10-424-599-224408	Sequence 224408, A
24	8	2.1	270	16	US-10-767-701-38282	Sequence 38282, A
25	8	2.1	366	15	US-10-335-977-7581	Sequence 7581, Ap
26	8	2.1	410	9	US-09-880-137-3	Sequence 3, Appl1
27	8	2.1	410	10	US-09-800-137A-3	Sequence 3, Appl1
28	8	2.1	418	9	US-09-880-137-1	Sequence 1, Appl1
29	8	2.1	418	9	US-09-880-137-2	Sequence 2, Appl1
30	8	2.1	418	10	US-09-800-137A-1	Sequence 1, Appl1
31	8	2.1	418	10	US-09-800-137A-2	Sequence 2, Appl1
32	8	2.1	418	14	US-10-043-487-348	Sequence 348, App
33	8	2.1	448	9	US-09-815-242-11562	Sequence 11562, A
34	8	2.1	448	15	US-10-282-122A-58890	Sequence 58890, A
35	8	2.1	448	15	US-10-335-977-7582	Sequence 7582, Ap
36	8	2.1	452	14	US-10-038-010-52	Sequence 52, Appl
37	8	2.1	574	16	US-10-437-963-172259	Sequence 172259, A
38	8	2.1	583	16	US-10-437-963-114326	Sequence 114326, A
39	8	2.1	593	15	US-10-425-114-56486	Sequence 56486, A
40	8	2.1	662	15	US-10-282-122A-67739	Sequence 67739, A
41	7	1.8	16	14	US-10-062-710-116	Sequence 116, App
42	7	1.8	16	14	US-10-239-313A-196	Sequence 196, App
43	7	1.8	19	15	US-10-446-234-4	Sequence 4, Appl1
44	7	1.8	57	15	US-10-424-599-200108	Sequence 200108, A
45	7	1.8	59	15	US-10-424-599-267779	Sequence 267779, A

ALIGNMENTS

RESULT 1
US-10-094-240-2
Sequence 2, Application US/10094240
Publication No. US20030082637A1
GENERAL INFORMATION:
APPLICANT: ZWIEBEL, LAURENCE J.
TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
FILE REFERENCE: N8289
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 10/056,405
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: 60/264,649
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 383
TYPE: PRT
ORGANISM: Anopheles gambiae
US-10-094-240-2
Query Match 100.0%; Score 383; DB 14; Length 383;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVYVFKYFKKCAPNGKVTLYMGKDPFDHYSGVEPIGIVYLVDEYTRDNKRVGQIYCS 60
DB 1 MVYVFKYFKKCAPNGKVTLYMGKDPFDHYSGVEPIGIVYLVDEYTRDNKRVGQIYCS 60
QY 61 FRYGRBDEVMGALNFOELCLASRQIYPRPEKSKQEQTKLOERLLKLGSAIIFFTNIS 120
DB 61 FRYGRBDEVMGALNFOELCLASRQIYPRPEKSKQEQTKLOERLLKLGSAIIFFTNIS 120
QY 121 PNAFSSVTLQGGEDNDGDPGVSVYVXIFAGESETRDTRHRSVTYLGIRKIQFAPYQOQ 180

```
Db 121 PNAASSVTLQOGEEDNGDPCGVSYVVKIFAGESEDTHTSRSTVTLGIRKIQFAPTKQDQ 180
QY 181 QPCTLVKDFMLSPGSELELVTLDKQYLHGERIGVNICIRNNSNMKVKIKAMWQGVND 240
Db 181 QPCTLVKDFMLSPGSELELVTLDKQYLHGERIGVNICIRNNSNMKVKIKAMWQGVND 240
QY 241 VVLFQNGSYRNTVASLETSEGCPIQPGSSLOKVMYTLPLSSNKKORRGIALDGOIKRQDQ 300
Db 241 VVLFQNGSYRNTVASLETSEGCPIQPGSSLOKVMYTLPLSSNKKORRGIALDGOIKRQDQ 300
QY 301 CLASTTLLAOPDORDAFVITISYAVKVLFGALGELSALPFLVLMHPKPGTKAKVIHA 360
Db 301 CLASTTLLAOPDORDAFVITISYAVKVLFGALGELSALPFLVLMHPKPGTKAKVIHA 360
QY 361 DSQADVETFRQDTIIDQASVDPE 383
Db 361 DSQADVETFRQDTIIDQASVDPE 383

RESULT 2
US-10-056-405-2
; Sequence 2, Application US/10056405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
; FILE REFERENCE: N7841
; CURRENT APPLICATION NUMBER: US/10/056,405
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-056-405-2

Query Match
Best Local Similarity 100.0%; Score 383; DB 14; Length 383;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYVNFVKFKCAMPNGKVTLYMGKRDVHVSGVEPIDGIIVLDEYIRDNKRVGQIVCS 60
Db 1 MYVNFVKFKCAMPNGKVTLYMGKRDVHVSGVEPIDGIIVLDEYIRDNKRVGQIVCS 60
QY 61 FRYGREDEVMGLNQKEICLASEQIYPRPEKSDKEOTKQERLTKLGSNAIPFTFNIS 120
Db 61 FRYGREDEVMGLNQKEICLASEQIYPRPEKSDKEOTKQERLTKLGSNAIPFTFNIS 120
QY 121 PNAASSVTLQOGEEDNGDPCGVSYVVKIFAGESEDTHTSRSTVTLGIRKIQFAPTKQDQ 180
Db 121 PNAASSVTLQOGEEDNGDPCGVSYVVKIFAGESEDTHTSRSTVTLGIRKIQFAPTKQDQ 180
QY 181 QPCTLVKDFMLSPGSELELVTLDKQYLHGERIGVNICIRNNSNMKVKIKAMWQGVND 240
Db 181 QPCTLVKDFMLSPGSELELVTLDKQYLHGERIGVNICIRNNSNMKVKIKAMWQGVND 240
QY 241 VVLFQNGSYRNTVASLETSEGCPIQPGSSLOKVMYTLPLSSNKKORRGIALDGOIKRQDQ 300
Db 241 VVLFQNGSYRNTVASLETSEGCPIQPGSSLOKVMYTLPLSSNKKORRGIALDGOIKRQDQ 300
QY 301 CLASTTLLAOPDORDAFVITISYAVKVLFGALGELSALPFLVLMHPKPGTKAKVIHA 360
Db 301 CLASTTLLAOPDORDAFVITISYAVKVLFGALGELSALPFLVLMHPKPGTKAKVIHA 360
QY 361 DSQADVETFRQDTIIDQASVDPE 383
Db 361 DSQADVETFRQDTIIDQASVDPE 383
```

```
RESULT 3
US-09-880-137-8
; Sequence 8, Application US/09880137
; Patent No. US20020031295A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: MNI-131
; CURRENT APPLICATION NUMBER: US/09/880,137
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-880-137-8

Query Match
Best Local Similarity 2.9%; Score 11; DB 9; Length 382;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELFPVLMHPRP 351
Db 339 ELFPVLMHPRP 349
```

```
RESULT 4
US-09-800-137A-8
; Sequence 8, Application US/09800137A
; Publication No. US20030157553A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: MNI-131
; CURRENT APPLICATION NUMBER: US/09/800,137A
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-800-137A-8

Query Match
Best Local Similarity 2.9%; Score 11; DB 10; Length 382;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELFPVLMHPRP 351
Db 339 ELFPVLMHPRP 349
```

```
RESULT 5
US-10-094-240-25
; Sequence 25, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N6289
; CURRENT APPLICATION NUMBER: US/10/094,240
```

```
/ CURRENT FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 10/056,405
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: 60/264,649
/ PRIOR FILING DATE: 2001-01-26
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 25
/ LENGTH: 398
/ TYPE: PRT
/ ORGANISM: Anopheles gambiae
US-10-094-240-25
```

```
Query Match          2.9%; Score 11; DB 14; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      62 RYGREDEVMG 72
      |||||
Db      61 RYGREDEVMG 71
```

```
RESULT 6
US-10-094-240-27
/ Sequence 27, Application US/10094240
/ Publication No. US20030082637A1
/ GENERAL INFORMATION:
/ APPLICANT: ZWIEREL, LAURENCE J.
/ TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
/ FILE REFERENCE: N8289
/ CURRENT APPLICATION NUMBER: US/10/094,240
/ CURRENT FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 10/056,405
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: 60/264,649
/ PRIOR FILING DATE: 2001-01-26
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 27
/ LENGTH: 401
/ TYPE: PRT
/ ORGANISM: Drosophila melanogaster
US-10-094-240-27
```

```
Query Match          2.9%; Score 11; DB 14; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      62 RYGREDEVMG 72
      |||||
Db      61 RYGREDEVMG 71
```

```
RESULT 7
US-09-880-137-4
/ Sequence 4, Application US/09880137
/ Patent No. US20020031295A1
/ GENERAL INFORMATION:
/ APPLICANT: Berstein, Gabriel
/ TITLE OF INVENTION: METHODS OF ASSAYING FOR G
/ TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
/ FILE REFERENCE: NMI-131
/ CURRENT APPLICATION NUMBER: US/09/880,137
/ CURRENT FILING DATE: 2001-03-05
/ PRIOR APPLICATION NUMBER: US 60/186,706
/ PRIOR FILING DATE: 2000-03-03
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 409
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-880-137-4
```

```
Query Match          2.9%; Score 11; DB 9; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      341 ELPEVLMHPRP 351
      |||||
Db      339 ELPEVLMHPRP 349
```

```
RESULT 8
US-09-880-137-7
/ Sequence 7, Application US/09880137
/ Patent No. US20020031295A1
/ GENERAL INFORMATION:
/ APPLICANT: Berstein, Gabriel
/ TITLE OF INVENTION: METHODS OF ASSAYING FOR G
/ FILE REFERENCE: NMI-131
/ CURRENT APPLICATION NUMBER: US/09/880,137
/ CURRENT FILING DATE: 2001-03-05
/ PRIOR APPLICATION NUMBER: US 60/186,706
/ PRIOR FILING DATE: 2000-03-03
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7
/ LENGTH: 409
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-880-137-7
```

```
Query Match          2.9%; Score 11; DB 9; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      341 ELPEVLMHPRP 351
      |||||
Db      339 ELPEVLMHPRP 349
```

```
RESULT 9
US-09-800-137A-4
/ Sequence 4, Application US/09800137A
/ Publication No. US20030157553A1
/ GENERAL INFORMATION:
/ APPLICANT: Berstein, Gabriel
/ TITLE OF INVENTION: METHODS OF ASSAYING FOR G
/ TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
/ FILE REFERENCE: NMI-131
/ CURRENT APPLICATION NUMBER: US/09/800,137A
/ CURRENT FILING DATE: 2001-03-05
/ PRIOR APPLICATION NUMBER: US 60/186,706
/ PRIOR FILING DATE: 2000-03-03
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 409
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-800-137A-4
```

```
Query Match          2.9%; Score 11; DB 10; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      341 ELPEVLMHPRP 351
      |||||
Db      339 ELPEVLMHPRP 349
```

RESULT 10

```

US-09-800-137A-7
; Sequence 7, Application US/09800137A
; Publication No. US20030157553A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: MNI-131
; CURRENT APPLICATION NUMBER: US/09/800,137A
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Homo sapiens mutation
US-09-800-137A-7

```

```

Query Match
Best Local Similarity 2.9%; Score 11; DB 10; Length 409;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 341 ELFPVLMHPPK 351
DB 339 ELFPVLMHPPK 349

```

```

RESULT 11
US-10-038-010-54
; Sequence 54, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: beta Arrestin2
; LOCATION: (1)..(410)
; OTHER INFORMATION:
US-10-038-010-54

```

```

Query Match
Best Local Similarity 2.9%; Score 11; DB 14; Length 410;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 341 ELFPVLMHPPK 351
DB 340 ELFPVLMHPPK 350

```

```

RESULT 12
US-10-722-357-31
; Sequence 31, Application US/10722357
; Publication No. US20040191803A1
; GENERAL INFORMATION:
; APPLICANT: GALLAGHER, MICHELA
; APPLICANT: LUND, PAULINE KAY
; APPLICANT: ROTHSTEIN, JEFFREY

```

```

; TITLE OF INVENTION: TARGET FOR THERAPY OF COGNITIVE IMPAIRMENT
; FILE REFERENCE: JHV-028.01
; CURRENT APPLICATION NUMBER: US/10/722,357
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: 60/413,152
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 31
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-722-357-31

```

```

Query Match
Best Local Similarity 2.9%; Score 11; DB 16; Length 466;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 341 ELFPVLMHPPK 351
DB 396 ELFPVLMHPPK 406

```

```

RESULT 13
US-09-880-137-5
; Sequence 5, Application US/09880137
; Patent No. US30020031295A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: MNI-131
; CURRENT APPLICATION NUMBER: US/09/880,137
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-137-5

```

```

Query Match
Best Local Similarity 2.3%; Score 9; DB 9; Length 388;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 22 GKRDVVDHV 30
DB 19 GKRDVVDHV 27

```

```

RESULT 14
US-09-880-137-6
; Sequence 6, Application US/09880137
; Patent No. US20020031295A1
; GENERAL INFORMATION:
; APPLICANT: Berstein, Gabriel
; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
; FILE REFERENCE: MNI-131
; CURRENT APPLICATION NUMBER: US/09/880,137
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/186,706
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-137-6

```

Query Match 2.3%: Score 9; DB 9; Length 388;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 GKRD FVDHV 30
 |||||
 Db 19 GKRD FVDHV 27

RESULT 15
 US-09-800-137A-5
 ; Sequence 5, Application US/09800137A
 ; Publication No. US20030157553A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berstein, Gabriel
 ; TITLE OF INVENTION: METHODS OF ASSAYING FOR G
 ; TITLE OF INVENTION: PROTEIN-COUPLED RECEPTOR LIGANDS AND MODULATORS
 ; FILE REFERENCE: WNI-131
 ; CURRENT APPLICATION NUMBER: US/09/800,137A
 ; CURRENT FILING DATE: 2001-03-05
 ; PRIOR APPLICATION NUMBER: US 60/166,706
 ; PRIOR FILING DATE: 2000-03-03
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 388
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-800-137A-5

Query Match 2.3%: Score 9; DB 10; Length 388;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22 GKRD FVDHV 30
 |||||
 Db 19 GKRD FVDHV 27

Search completed: February 10, 2005, 00:13:29
 Job time : 132 secs

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 9, 2005, 23:46:08 ; Search time 176 Seconds
(without alignments)
1114.354 Million cell updates/sec

Title: US-10-056-405-2
Perfect score: 383
Sequence: 1 MVNFKVFKKCAPNGKVTLY.....ADVETFRQDTIDQASVDFE 383

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Uniprot 03:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	383	100.0	383	2 Q95NF3	Q95nf3 anopheles g
2	383	100.0	417	2 Q7PMG5	Q7pmg5 anopheles g
3	245	64.0	245	2 Q6VPP0	Q6vpp0 anopheles g
4	225	6.5	381	2 Q9BIC9	Q9bic9 acaulapnus
5	20	5.2	364	1 ARRA_DROME	P15372 drosophila
6	18	4.7	18	2 Q7YTV9	Q7ytw9 anopheles s
7	18	4.7	18	2 Q7YTW0	Q7ytw0 anopheles p
8	18	4.7	18	2 Q7YTW1	Q7ytw1 anopheles a
9	17	4.4	363	1 ARRI_CALVI	P51486 calliphora
10	16	4.2	381	1 ARRH_HELV1	P55274 heliothis v
11	13	3.7	52	2 Q6X126	Q6x126 drosophila
12	13	3.4	400	1 ARRH_LIMPO	P14844 limulus pol
13	13	3.1	407	1 ARRH_LOCM1	P32122 locusta mig
14	12	2.9	217	2 Q68DZ5	Q68daz5 homo sapien
15	11	2.9	398	2 Q66GUE	Q66gue anopheles g
16	11	2.9	400	1 ARRB_CALVI	P51487 calliphora
17	11	2.9	401	1 ARRB_DROME	P19107 drosophila
18	11	2.9	401	1 ARRB_DROMI	P19108 drosophila
19	11	2.9	401	2 Q705Q8	Q705q8 anopheles g
20	11	2.9	405	1 ARR2_ONCMY	P51467 oncorhynchu
21	11	2.9	406	2 Q6PFF2	Q6pff2 brachydanio
22	11	2.9	407	1 ARRI_ONCMY	P51466 oncorhynchu
23	11	2.9	408	2 Q6DFC4	Q6dfc4 xenopus lae
24	11	2.9	408	2 Q6GPy2	Q6gpy2 xenopus lae
25	11	2.9	409	1 Q7T2D2	Q7t2d2 brachydanio
26	11	2.9	409	1 ARR2_HUMAN	P32121 homo sapien
27	11	2.9	409	2 Q6ICT3	Q6ict3 homo sapien
28	11	2.9	410	1 ARR2_MOUSE	Q91914 mus musculu
29	11	2.9	410	1 ARR2_RAT	P29067 rattus norv
30	11	2.9	415	1 ARR3_ONCMY	P51468 oncorhynchu
31	11	2.9	420	1 ARR2_BOVIN	P32120 bos taurus

32	10	2.6	147	2 Q6WGR2	Q6wgr2 metapenaeus
33	9	2.3	359	2 Q96EN2	Q96en2 homo sapien
34	9	2.3	388	1 ARRC_HUMAN	P16575 homo sapien
35	9	2.3	392	2 Q9N0H5	Q9n0h5 bos taurus
36	9	2.3	392	2 Q6NVY2	Q6nv2 brachydanio
37	9	2.3	392	2 Q9PTE7	Q9pce7 ambystoma t
38	9	2.3	394	2 Q7Y5T8	Q7y5t8 sus scrofa
39	9	2.3	412	2 Q64ID8	Q64id8 xenopus lae
40	9	2.3	470	2 Q9V393	Q9v393 drosophila
41	9	2.3	1773	2 Q82745	Q82745 arabidopsis
42	8	2.1	141	2 Q48240	Q48240 dennisus dis
43	8	2.1	142	2 Q48239	Q48239 dennisus dis
44	8	2.1	143	2 Q48243	Q48243 dennisus car
45	8	2.1	147	2 Q48246	Q48246 dennisus sin

ALIGNMENTS

RESULT 1

ID	Q95NF3	PRELIMINARY:	PRT:	383 AA.
AC	Q95NF3			
DT	01-DEC-2001 (TREMBlrel. 19, Created)			
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)			
DT	05-JUL-2004 (TREMBlrel. 27, Last annotation update)			
DE	Arrestin.			
GN	Name-Arri;			
OS	Anopheles gambiae (African malaria mosquito).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.			
OX	NCBI_TaxId=7165;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Antennae, and Head;			
RX	MEDLINE=21680430; PubMed=11822731;			
RA	Ricci I., Santolamazza F., Costantini C., Favia G.;			
RT	"Molecular characterization and chromosomal mapping of transcripts			
RT	having tissue-specific expression in the malaria mosquito anopheles			
RT	gambiae: possible involvement in visual or olfactory processes.";			
RL	Parasitol. Res. 88:1-8(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=G3;			
RX	MEDLINE=21819361; PubMed=11792843; DOI=10.1073/pnas.025054999;			
RA	Merrill C.E., Riesgo-Escovar J., Plets R.J., Kafatos F.C.;			
RA	Carlson J.R., Zwiebel L.U.;			
RT	"Visual arrestin in olfactory pathways of Drosophila and the malaria			
RT	vector mosquito Anopheles gambiae.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:1633-1638(2002).			
DR	EMBL; AJ304409; CAC39103.2; -.			
DR	EMBL; AY017417; AAC54081.1; -.			
DR	HSSP; P1870; IG4M.			
DR	GO; GO:007600; P:sensory perception; IEA.			
DR	GO; GO:0007165; P:signal transduction; IEA.			
DR	InterPro; IPR000698; Arrestin.			
DR	InterPro; IPR011022; Arrestin_C.			
DR	InterPro; IPR011021; Arrestin_N.			
DR	Pfam; PF02752; Arrestin_C_1.			
DR	Pfam; PF00339; Arrestin_N_1.			
DR	PRINTS; PR00309; ARRESTIN.			
DR	PROSITE; PS00295; ARRESTIN; 1.			
DR	SEQUENCE 383 AA; 42809 MW; FA946438592B53E CRC64;			
Qy	Query Match	100.0%;	Score 383;	DB 2; Length 383;
Qy	Best Local Similarity	100.0%;	Pred. No. 0;	
Qy	Matches 383;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1 MVNFKVFKKCAPNGKVTLYMGKRDVDFHSGVEPIDGIIVLDEYIRDNKVGQIVCS 60			
Qy	1 MVNFKVFKKCAPNGKVTLYMGKRDVDFHSGVEPIDGIIVLDEYIRDNKVGQIVCS 60			
Qy	61 FRVGRDEVMGLNFQKELCLASQIYPRPEKSDKEQTKLQERLLKLKLSNAIPFTNIS 120			

Db 61 FRFGREDEVMGINFQKELCLASEQIYRPEKSDKEQTKQERLLKKGNAIPFTFNIS 120
Qy 121 PNAPSSVTLQOGEDDNGPCGVSYVVKIFAGESETDTRHRSVTTLGIRKIOPAPTKQOQ 180
Db 121 PNAPSSVTLQOGEDDNGPCGVSYVVKIFAGESETDTRHRSVTTLGIRKIOPAPTKQOQ 180
Qy 181 OPTCLVRKDFMLSPGELLEVTLDKQYLHGERIGVNICIRNNSNMVKKIKAMVQOQVD 240
Db 181 OPTCLVRKDFMLSPGELLEVTLDKQYLHGERIGVNICIRNNSNMVKKIKAMVQOQVD 240
Qy 241 VVLFONGSYRNTVASLFTSECCPIQPGSSIQKVMYLTPLSSNKQRRGIALDGOIKRQDQ 300
Db 241 VVLFONGSYRNTVASLFTSECCPIQPGSSIQKVMYLTPLSSNKQRRGIALDGOIKRQDQ 300
Qy 301 CLASTTLAOPDQDADFVGIISYAVVKVFLGALGSELAEPLVLMHPKGTAKVYIHA 360
Db 301 CLASTTLAOPDQDADFVGIISYAVVKVFLGALGSELAEPLVLMHPKGTAKVYIHA 360
Qy 361 DSQADVETFRQDITIDQASVDPE 383
Db 361 DSQADVETFRQDITIDQASVDPE 383

RESULT 2

Q7PMG5 PRELIMINARY; PRT; 417 AA.
ID 07PMG5
AC 07PMG5;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE ENSANGP0000012569 (Fragment).
GN Name=ENSANGG0000010080;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxId=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
DR EMBL; AAAB01008980; EAA13874.2; -
DR HSSP; P17870; IGAM.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR011022; Arrestin_C.
DR InterPro; IPR011021; Arrestin_N.
DR Pfam; PF02752; Arrestin_C; 1.
DR Pfam; PF00339; Arrestin_N; 1.
DR PRINTS; PR00309; ARRESTIN.
DR PRODOM; PD002099; ARRESTIN.
DR PROSITE; PS00295; ARRESTINS; 1.
FT NON TER 1
SQ SEQUENCE 417 AA; 46591 MW; 4EB98A3C32257FC0 CRC64;

Query Match

Best Local Similarity 100.0%; Score 383; DB 2; Length 417;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYNFKVFKKCAKNGVTLVYMGKRPVHVHSGVEPIDIGIVLDDIYLRDNKRVGOIVCS 60
Db 35 MYNFKVFKKCAKNGVTLVYMGKRPVHVHSGVEPIDIGIVLDDIYLRDNKRVGOIVCS 60
Qy 61 FRFGREDEVMGINFQKELCLASEQIYRPEKSDKEQTKQERLLKKGNAIPFTFNIS 120
Db 95 FRFGREDEVMGINFQKELCLASEQIYRPEKSDKEQTKQERLLKKGNAIPFTFNIS 154
Qy 121 PNAPSSVTLQOGEDDNGPCGVSYVVKIFAGESETDTRHRSVTTLGIRKIOPAPTKQOQ 180

Db 155 PNAPSSVTLQOGEDDNGPCGVSYVVKIFAGESETDTRHRSVTTLGIRKIOPAPTKQOQ 214
Qy 181 OPTCLVRKDFMLSPGELLEVTLDKQYLHGERIGVNICIRNNSNMVKKIKAMVQOQVD 240
Db 215 OPTCLVRKDFMLSPGELLEVTLDKQYLHGERIGVNICIRNNSNMVKKIKAMVQOQVD 240
Qy 241 VVLFONGSYRNTVASLFTSECCPIQPGSSIQKVMYLTPLSSNKQRRGIALDGOIKRQDQ 300
Db 241 VVLFONGSYRNTVASLFTSECCPIQPGSSIQKVMYLTPLSSNKQRRGIALDGOIKRQDQ 300
Qy 301 CLASTTLAOPDQDADFVGIISYAVVKVFLGALGSELAEPLVLMHPKGTAKVYIHA 360
Db 335 CLASTTLAOPDQDADFVGIISYAVVKVFLGALGSELAEPLVLMHPKGTAKVYIHA 394
Qy 361 DSQADVETFRQDITIDQASVDPE 383
Db 395 DSQADVETFRQDITIDQASVDPE 417

RESULT 3

Q6VPP0 PRELIMINARY; PRT; 245 AA.
ID Q6VPP0
AC Q6VPP0;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Arrestin (Fragment).
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxId=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4arr, Yaounde, and I3-5;
RA Morlais I., Poncon N., Simard F.
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY333991; AAR01116.1; -
DR EMBL; AY333992; AAR01117.1; -
DR EMBL; AY333993; AAR01118.1; -
DR EMBL; AY333995; AAR01120.1; -
DR EMBL; AY333990; AAR01115.1; -
DR EMBL; AY333996; AAR01121.1; -
DR EMBL; AY333994; AAR01119.1; -
DR HSSP; P08168; IAYR.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR011022; Arrestin_C.
DR InterPro; IPR011021; Arrestin_N.
DR Pfam; PF02752; Arrestin_C; 1.
DR Pfam; PF00339; Arrestin_N; 1.
DR PRINTS; PR00309; ARRESTIN.
DR PRODOM; PD002099; ARRESTIN.
FT NON TER 1
SQ SEQUENCE 245 AA; 27038 MW; 86D8F4520115DA5F CRC64;

Query Match

Best Local Similarity 64.0%; Score 245; DB 2; Length 245;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 LQOGEDDNGPCGVSYVVKIFAGESETDTRHRSVTTLGIRKIOPAPTKQOQCTLVK 188
Db 1 LQOGEDDNGPCGVSYVVKIFAGESETDTRHRSVTTLGIRKIOPAPTKQOQCTLVK 188
Qy 189 DFMLSPGELLEVTLDKQYLHGERIGVNICIRNNSNMVKKIKAMVQOQVDVLFONGS 248
Db 61 DFMLSPGELLEVTLDKQYLHGERIGVNICIRNNSNMVKKIKAMVQOQVDVLFONGS 120
Qy 249 YRNTVASLFTSECCPIQPGSSIQKVMYLTPLSSNKQRRGIALDGOIKRQDCLASTTL 308
Db 121 YRNTVASLFTSECCPIQPGSSIQKVMYLTPLSSNKQRRGIALDGOIKRQDCLASTTL 180

```

Qy 309 ACPDORDAFVITISYAVKVLFLGALGELSAELPFVLMHPKPTKAKVTHADSOADVERT 368
    |||||
Db 181 ACPDORDAFVITISYAVKVLFLGALGELSAELPFVLMHPKPTKAKVTHADSOADVERT 240
    |||||

Qy 369 PRODT 373
    |||||
Db 241 PRODT 245

RESULT 4
ID C9B1G9 PRELIMINARY; PRT; 381 AA.
C9B1G9
AC C9B1G9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Arrectin.
GN Name=arcti;
OS Acalaphus macaronius.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Neuroptera; Acalaphidae; Acalaphus.
OX NCBI_TaxID=146496;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=21184354; PubMed=11287006; DOI=10.1016/S0014-5793(01)02287-6;
RA Bentrop J., Schillo M., Gerdon G., Draetler G., Paulsen R.;
RT "UV-light-dependent binding of a visual arrestin 1 isoform to
RT photoreceptor membranes in a neuropteran (Acalaphus) compound eye.";
RL FEBS Lett. 493:112-116(2001).
DR EMBL; AJ303080; CAC36938.1; -.
DR HSSP; P17870; 1G4M.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000698; Arrectin.
DR InterPro; IPR011022; Arrectin_C.
DR InterPro; IPR011021; Arrectin_N.
DR Pfam; PF02752; Arrectin_C_1.
DR Pfam; PF00339; Arrectin_N_1.
DR PRINTS; PR00309; ARRECTIN.
DR PROSITE; PS003295; ARRECTIN.
SQ SEQUENCE 381 AA; 42789 MW; 451DEF465969857D CRC64;

Query Match 6.5%; Score 25; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 5,9e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 RKVFGQIVCSFRYGRDEWVGLNF 75
    |||||
Db 52 RKVFGQIVCSFRYGRDEWVGLNF 76
    |||||

RESULT 5
ARRA_DROME STANDARD; PRT; 364 AA.
AC P15373; O9VJ28;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Phoresstin II (Arrectin A) (Arrectin 1).
GN Name=arcti; Synonym=ARRA; ORFNames=CG5711;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
RX MEDLINE=90138925; PubMed=168956;
RA Smith D.P., Sheih B.-H., Zuker C.S.;
RT "Isolation and structure of an arrestin gene from Drosophila.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1003-1007(1990).

```

```

RN 121
RP SEQUENCE FROM N.A.
RX MEDLINE=90138926; PubMed=2105491;
RA Hyde D.R., Mecklenburg K.L., Pollock J.A., Vithelec T.S., Benzer S.;
RT "Twenty Drosophila visual system cDNA clones: one is a homolog of
RT human arrestin.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1008-1012(1990).
RN 131
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goodyne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt A., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayani A., An H.-U., Andrews-Planckoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brodter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey S., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ilegwam C.,
RA Jitali M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskearn D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Styrekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN 141
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitefield E.J., Bayraktaroglu L., Berman B.P.,
RA Belencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.B.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN 151
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Head;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kromliller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN 161
RP PHOSPHORYLATION.

```

```

RX MEDLINE=91282780; PubMed=1905538;
RA Matsumoto H., Yamada T.;
RT "Phosretin I and II: arrestin homologs which undergo differential
RT light-induced phosphorylation in the Drosophila photoreceptor in
RL vivo."
RN Biochem. Biophys. Res. Commun. 177:1306-1312 (1991).
RP (7)
RX FUNCTION.
RA MEDLINE=93303590; PubMed=8316831;
RA Dolph P.J., Ranganathan R., Colley N.J., Hardy R.W., Socolich M.,
RA Zuker C.S.;
RT "Arrestin function in inactivation of G protein-coupled receptor
RT rhodopsin in vivo."
RL Science 260:1910-1916 (1993).
CC -1- FUNCTION. Regulates photoreceptor cell deactivation. Arr1 and Arr2
CC proteins are mediators of rhodopsin inactivation and are essential
CC for the termination of the phototransduction cascade.
CC -1- TISSUE SPECIFICITY. Expressed specifically and abundantly in the
CC photoreceptors. Inner and outer segments, and the inner plexiform
CC regions of the retina.
CC -1- PTM: Phosphorylated, but does not undergo light-induced
CC phosphorylation.
CC -1- SIMILARITY: Belongs to the arrestin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30177; -; NOT_ANNOTATED_CDS.
DR EMBL; M30140; AAA28380.1; -
DR EMBL; AE003657; AAF53644.1; -
DR EMBL; AY061824; AAL27635.1; -
DR PIR; A34867; A34867.
DR HSSP; P17870; 1G4M.
DR INTCAT; P15372; -
DR FLYBASE; FBgn0000120; Arr1.
DR GO; GO:0005624; C:membrane fraction; IDA.
DR GO; GO:0016028; C:rhodome; IDA.
DR GO; GO:0016060; P:metarhodopsin inactivation; IGI.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR011022; Arrestin_C.
DR InterPro; IPR011021; Arrestin_N.
DR Pfam; PF02752; Arrestin_N.
DR Pfam; PF00339; Arrestin_C.
DR PRINTS; PR00309; ARRESTIN.
DR PRODOM; PD002039; Arrestin; 2.
DR PROSITE; PS00295; ARRESTINS; 1.
DR Phosphorylation; Sensory transduction; Vision.
KW SEQUENCE 364 AA; 40771 MW; 0DCCT64CF890FC2 CRC64;
SQ
Query Match 5.2%; Score 20; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 5.8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 228 VKKIKAMVQGVDFVLFONG 247
DB 227 VKKIKAMVQGVDFVLFONG 246

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=30659;
RN [1]
RP SEQUENCE FROM N.A.
RA Favia G., Ricci I., Castagnoli M., Esposito F.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ544226; CAD66651.1; -
FT NON_TER 18
SQ SEQUENCE 18 AA; 2075 MW; A273A394A2A6EF66 CRC64;
Query Match 4.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVYNFKVFKKCAPNGKVT 18
DB 1 MVYNFKVFKKCAPNGKVT 18

```

RESULT 7

```

QYTW0
ID QYTW0 PRELIMINARY; PRT; 18 AA.
AC QYTW0;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE Arrestin (Fragment).
GN Name=arr1;
OS Anopheles pharensis.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=221566;
RN [1]
RP SEQUENCE FROM N.A.
RA Favia G., Ricci I., Castagnoli M., Esposito F.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ544225; CAD66650.1; -
FT NON_TER 18
SQ SEQUENCE 18 AA; 2075 MW; A273A394A2A6EF66 CRC64;
Query Match 4.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVYNFKVFKKCAPNGKVT 18
DB 1 MVYNFKVFKKCAPNGKVT 18

```

RESULT 8

```

QYTW1
ID QYTW1 PRELIMINARY; PRT; 18 AA.
AC QYTW1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE Arrestin (Fragment).
GN Name=arr1;
OS Anopheles arabiensis (Mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=7173;
RN [1]
RP SEQUENCE FROM N.A.
RA Favia G., Ricci I., Castagnoli M., Esposito F.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ544224; CAD66649.1; -
FT NON_TER 18
SQ SEQUENCE 18 AA; 2075 MW; A273A394A2A6EF66 CRC64;
Query Match 4.7%; Score 18; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVYNFKVFKKCAPNGKVT 18
DB 1 MVYNFKVFKKCAPNGKVT 18

```

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVTNFKVFKKCAPNGKYT 18
 Db 1 MVTNFKVFKKCAPNGKYT 18

RESULT 9

ARR1_CALVI STANDARD; PRT; 363 AA.

AC P51466;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Phosrestin II (Arrestin A) (Arrestin 1).
 GN Name-ARR1;
 OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
 CC Calliphoridae; Calliphora.
 NCBI_TaxID=7373;
 RX MEDLINE=95014564; PubMed=7929436;
 RA Plangger A., Malicki D., Whitney M., Paulsen R.;
 RT "Mechanism of arrestin 2 function in rhodometric photoreceptors."
 RL J. Biol. Chem. 269:26969-26975(1994).
 CC -1- SIMILARITY: Belongs to the arrestin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X79072; CAA55672.1; -.
 DR PIR; A55081; A55081.
 DR HSSP; P17870; 1G4M.
 DR InterPro; IPR000698; Arrestin.
 DR InterPro; IPR011022; Arrestin_C.
 DR InterPro; IPR011021; Arrestin_N.
 DR Pfam; PF02752; Arrestin_C_1.
 DR Pfam; PF00339; Arrestin_N_1.
 DR PRINTS; PR00309; ARRESTIN.
 DR ProDom; PD002099; Arrestin; 2.
 DR PROSITE; PS00295; ARRESTINS; 1.
 KM Sensory transduction; Vision.
 SQ SEQUENCE 363 AA; 40680 MW; 7C345D818E46C23E CRC64;

Query Match 4.4%; Score 17; DB 1; Length 363;
 Best Local Similarity 100.0%; Pred. No. 5.9e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 SYAVKVLFLGALGCEL 338
 Db 321 SYAVKVLFLGALGCEL 337

RESULT 10

ARRH_HELVI STANDARD; PRT; 381 AA.

AC P55274;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Arrestin homolog.
 OS Heliothis virescens (Noctuid moth) (Owllet moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
 OC Noctuidae; Heliothinae; Heliothis.

OX NCBI_TaxID=7102;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Antenna;

RX MEDLINE=93199955; PubMed=8452755; DOI=10.1016/0898-6568(93)90009-B;

RA Raming K., Freitag J., Krieger J., Brier H.;

RT "Arrestin-subtypes in insect antennae."

RL Cell. Signal. 5:69-80(1993).

CC -1- SIMILARITY: Belongs to the arrestin family.

DR PIR; B56607; B56607.

DR HSSP; P17870; 1G4M.

DR InterPro; IPR000698; Arrestin.

DR InterPro; IPR011022; Arrestin_C.

DR InterPro; IPR011021; Arrestin_N.

DR Pfam; PF02752; Arrestin_C_1.

DR Pfam; PF00339; Arrestin_N_1.

DR PRINTS; PR00309; ARRESTIN.

DR ProDom; PD002099; Arrestin; 2.

DR PROSITE; PS00295; ARRESTINS; 1.

KM Sensory transduction.

SQ SEQUENCE 381 AA; 42747 MW; 84BB92B1B3DA573 CRC64;

Query Match 4.2%; Score 16; DB 1; Length 381;
 Best Local Similarity 100.0%; Pred. No. 6.2e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVTNFKVFKKCAPNGK 16
 Db 1 MVTNFKVFKKCAPNGK 16

RESULT 11

06X126 PRELIMINARY; PRT; 52 AA.

AC 06X126;

DT 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

DE Similar to Drosophila melanogaster CG5711 (Fragment).

OS Drosophila yakuba (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

CC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7245;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22887302; PubMed=14525923; DOI=10.1101/gr.1311003;

RA Domazet-Lošo T., Tautz D.;

RT "An evolutionary analysis of orphan genes in Drosophila."

RL Genome Res. 13:2213-2219(2003).

DR EMBL; AY232006; AAR10029.1; -.

DR HSSP; P08168; 1AYR.

DR GO; GO:0007600; P:sensory perception; IEA.

DR GO; GO:0007165; P:signal transduction; IEA.

DR InterPro; IPR000698; Arrestin.

DR ProDom; PD002099; Arrestin; 1.

FT NON_TER 1 52

FT NON_TER 1 52

SQ SEQUENCE 52 AA; 6166 MW; 081C148570B5EB6F CRC64;

Query Match 3.7%; Score 14; DB 2; Length 52;
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 VEPIDGIVVDDEY 46
 Db 33 VEPIDGIVVDDEY 46

RESULT 12

ARRH_LIMPO STANDARD; PRT; 400 AA.

AC P51484;

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 15-JUL-1999 (Rel. 38, last annotation update)
DE Arrestin, lateral eye
OS Limulus polyphemus (Atlantic horseshoe crab)
CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OX NCB1_Taxid=6850;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA MEDLINE=95096819; PubMed=7798902;
RA Smith W.C., Greenberg R.M., Calman B.G., Hendrix M.M., Hutchinson L.,
RT Donoso L.A., Bettle B.-A.;
RT "Isolation and expression of an arrestin cDNA from the horseshoe crab
RT lateral eye.";
RL J. Neurochem. 64:1-13(1995).
CC -1- FUNCTION: Plays an important role in the photoreceptor
CC transduction.
CC -1- PTM: Phosphorylated.
CC -1- SIMILARITY: Belongs to the arrestin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08883; AAA82007.1; -.
DR HSSP; P17870; 1G4M.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR011022; Arrestin_C.
DR InterPro; IPR011021; Arrestin_N.
DR Pfam; PF02752; Arrestin_C_1.
DR Pfam; PF00339; Arrestin_N_1.
DR PRINTS; PR00309; ARRESTIN.
DR ProDom; PD002099; Arrestin; 2.
DR PROSITE; PS00295; ARRESTINS; 1.
DR Phosphorylation; Sensory transduction; Vision.
SQ SEQUENCE 400 AA; 44302 MW; F3DD0D25BC2BEE33 CRC64;

Query Match
Best local Similarity 3.4%; Score 13; DB 1; Length 400;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 FRYGREDEVMGL 73
DB 68 FRYGREDEVMGL 80

RESULT 13
ARRH_LOCMT STANDARD; PRT; 407 AA.
AC P32122;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, last sequence update)
DE 15-JUL-1999 (Rel. 38, last annotation update)
DE Arrestin homolog.
OS Locusta migratoria (Migratory locust).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
CC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCB1_Taxid=7004;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Antenna;
RA MEDLINE=9319955; PubMed=8452755; DOI=10.1016/0898-6568(93)90009-B;
RT Ramling K., Freitag J., Krieger J., Breer H.;
RT "Arrestin-subtypes in insect antennae.";
RL Cell. Signal. 5:69-80(1993).
CC -1- SIMILARITY: Belongs to the arrestin family.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S57174; AAB25860.1; -.
DR PIR; A56607; A56607.
DR HSSP; P17870; 1G4M.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR011022; Arrestin_C.
DR InterPro; IPR011021; Arrestin_N.
DR Pfam; PF02752; Arrestin_C_1.
DR Pfam; PF00339; Arrestin_N_1.
DR PRINTS; PR00309; ARRESTIN.
DR ProDom; PD002099; Arrestin; 2.
DR PROSITE; PS00295; ARRESTINS; 1.
DR Sensory transduction.
SQ SEQUENCE 407 AA; 45543 MW; DEC28A3A534935BB CRC64;

Query Match
Best local Similarity 3.1%; Score 12; DB 1; Length 407;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 RYGREDEVMGL 73
DB 69 RYGREDEVMGL 80

RESULT 14
Q68DZ5 PRELIMINARY; PRT; 217 AA.
AC Q68DZ5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, last annotation update)
DE Hypothetical protein DKFZp686L0365.
GN Name=DKFZp686L0365;
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endometrium;
RA The German cDNA Consortium;
RA Ottenwälder B., Obermaier B., Deuschendorf S., Schaipe A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Pöbo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBS databases.
DR EMBL; CR749218; CAH18075.1; -.
DR InterPro; IPR000698; Arrestin.
DR InterPro; IPR011022; Arrestin_C.
DR Pfam; PF02752; Arrestin_C_1.
DR ProDom; PD002099; Arrestin; 1.
DR Hypothetical protein.
SQ SEQUENCE 217 AA; 24362 MW; 61653F5BACBB0FD CRC64;

Query Match
Best local Similarity 2.9%; Score 11; DB 2; Length 217;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 ELPFVLMHPKP 351
DB 147 ELPFVLMHPKP 157

RESULT 15
Q66GU6 PRELIMINARY; PRT; 398 AA.
AC Q66GU6;

```

```
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Arreclin 2-like protein Ariz2.
GN Name=ARR2;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RX MEDLINE=23159591; Pubmed=14986925;
RA Merril C.E., Pitsa R.J., Zwiebel L.J.;
RT "Molecular characterization of arresclin family members in the malaria
RT vector mosquito, Anopheles gambiae.";
RL Insect Mol. Biol. 12:641-650(2003).
DR EMBL; BK000996; DAA00888.1; -.
DR InterPro; IPR000698; Arresclin.
DR InterPro; IPR011022; Arresclin_C.
DR InterPro; IPR011021; Arresclin_N.
DR Pfam; PF02752; Arresclin_C; 1.
DR Pfam; PF00339; Arresclin_N; 1.
DR PRINTS; PR00309; ARRESTIN.
DR PRODOM; PD002099; Arresclin; 2.
DR PROSITE; PS00295; ARRESTINS; 1.
SQ SEQUENCE 398 AA; 44536 MW; 73DD73FB34F01418 CRC64;
```

Query Match 2.9%; Score 11; DB 2; Length 398;
Best Local Similarity 100.0%; Pred.No. 0.066;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 62 RYGRDEDEVNG 72
   |||||
Db 61 RYGRDEDEVNG 71
```

Search completed: February 10, 2005, 00:01:31
Job time : 177 secs

This Page Blank (nspto)

This Page Blank (nspto)